NICHOLSON 1963

AEC RESEARCH AND DEVELOPMENT REPORT

HW-76279

FIXED TIME ESTIMATION OF COUNTING RATES WITH BACKGROUND CORRECTIONS

W. L. NICHOLSON

SEPTEMBER 1963

HANFORD LABORATORIES

HANFORD ATOMIC PRODUCTS OPERATION
RICHLAND, WASHINGTON

GENERAL ELECTRIC

HW-76279

UC-32, Mathematics
and Computers
(TID-4500, 24th Ed.)

FIXED TIME ESTIMATION OF COUNTING RATES WITH BACKGROUND CORRECTIONS

By

W. L. Nicholson

Applied Mathematics Operation
Hanford Laboratories

September 1963

HANFORD ATOMIC PRODUCTS OPERATION
RICHLAND, WASHINGTON

Work performed under Contract No. AT(45-1)-1350 between the Atomic Energy Commission and General Electric Company

Printed by/for the U. S. Atomic Energy Commission

Printed in USA. Price $ 3.00 Available from the
Office of Technical Services
Department of Commerce
Washington 25, D.C.

TABLE OF CONTENTS

FIXED TIME ESTIMATION OF COUNTING RATES WITH BACKGROUND CORRECTIONS

I. INTRODUCTION

A routine method for estimating the total radioactivity in a long-lived sample involves monitoring the sample with a detector for a fixed time period t, recording the total number of counts x, and calculating a counting rate x/t. A background counting rate estimate y/s is also calculated for the same detector from a background count y recorded over a fixed time period s. The background estimate, which supposedly represents all spurious activity, is subtracted to give the sample net counting rate estimate

$$\tilde{c} = x/t - y/s. \tag{1}$$

The interpretation of the estimate $\tilde{c}$ must now be considered. Let C be the true net counting rate of the sample and let B be the true background counting rate during the period over which the sample was counted. So, x/t is an estimate of C + B, and y/s is an estimate of B. Within this framework several questions are pertinent. The first of these is

Q_1. Does the estimate $\tilde{c}$ imply that $C>0$; i.e., does the net sample counting rate imply the existence of "something" in the sample in addition to background activity?

An answer of "NO" to question Q_1 actually is a decision that C is less than some nominal threshold C_0. The second of these logically follows:

Q_2. What is the threshold C_0; i.e., what is the minimum C which can be detected as positive a reasonable portion of the time? Here, "reasonable" is open to definition by the economics of the situation.

An answer of "YES" to question Q_1 leads to the refinement:

Q_3. How good an estimate of C is $\tilde{c}$, and are there others that are better in some sense?

This paper considers the answers to the above questions and related topics for the case of appreciable background. In terms of expected total net sample count tC and expected total background count tB in the sample gross count, appreciable background means

$$0 \leqq (tC)/(tB)^{1/2} < 15.$$

The specific recommendations in the form of hypotheses testing and confidence interval estimation rules for answering questions Q_1, Q_2, and Q_3 are stated in Section II. A short discussion of the statistical concepts used in the paper is contained in Section III. Those interested only in the application and interpretation of the rules should read Sections I, II, and possibly III, plus referenced appendix material. The investigations supporting these recommendations are the subjects of Sections IV, V, and VI plus additional appendix material.

To answer the above questions some information about the distributional properties of x and y must be available. In the sequel we assume that the two counts x and y are random observations on Poisson distributions* with mean value parameters (C + B)t and Bs, respectively. With these assumptions, x/t is an unbiased estimate of C + B with variance (C + B)/t, and y/s is an unbiased estimate of B with variance B/s.

Questions Q_1 and Q_2 can be answered within the framework of classical statistical theory by setting up the null hypothesis H_o that C = 0

---

* The assumption that x is a Poisson distribution random quantity imposes the restriction that the total counting period t is much less than the half-life $t_{1/2}$ of the sample. If this condition is not satisfied, the Poisson variance (C + B)t appreciably overestimates the actual variance of x, which in turn, leads to unnecessarily broad confidence statements. For example, with t = 0.19 $t_{1/2}$, confidence intervals based on the Poisson distribution are about 10% wider than the correct ones, and with t = 0.43 $t_{1/2}$ they are about 25% wider than the correct ones.

and testing it against the composite alternative hypothesis H_C that $C > 0$. For any decision rule D for deciding between H_o and H_C, the power function provides the answers to Q_1, and Q_2. In Section IV, four decision rules D_1, D_2, D_3, and D_e are compared in detail. The first three of these rules are:

D_1. Decide H_C is true if $\tilde{c} > n_\alpha \, [x/t^2 + y/s^2]^{1/2}$;

Otherwise, decide H_o is true.

D_2. Decide H_C is true if $\tilde{c} > n_\alpha \, [(1/t + 1/s)(y/s)]^{1/2}$; (2)

Otherwise, decide H_o is true.

D_3. Decide H_C is true if $\tilde{c} > n_\alpha \, [(x+y)/(st)]^{1/2}$;

Otherwise, decide H_o is true.

Here, n_α is the $100(1-\alpha)$ percentile of the unit normal distribution. Customarily an investigator uses a 1%, 5%, or 10% level of significance decision rule to test the null hypothesis H_o. The corresponding n_α values are $n_{0.01} = 2.326$, $n_{0.05} = 1.645$, and $n_{0.10} = 1.282$, respectively.

All three of the rules in (2) are based upon asymptotic theory; specifically, that for large expected value in the Poisson distribution is approximately normal. In the limit as $(C + B)t$ and Bs approach infinity, $\tilde{c}$ divided by the bracketed quantity on the right of each inequality in (2) has a unit normal distribution, so the rules all have exactly $100\alpha\%$ significance levels. In Section IV the behavior of the rules is investigated for finite expected values. Previous observations of a number of authors, e.g.[1, 2], concerning the goodness of the normal approximation to the distribution of the difference of the two Poisson variables for small expected values are substantiated. This goodness carries over to the distribution of the decision rules in (2).

Rules D_1 and D_2 are common methods of answering questions Q_1 and Q_2 familiar to users of counting instruments. D_3, while probably not so familiar, turns out to be more robust than the others. Power surface calculations show that the stated D_3 level of significance is closest to the true one over a reasonable range of B, t and s values.

The differences in D_1, D_2, and D_3 are in the handling of the variance estimate of $\tilde{c}$. D_1 uses the obvious unbiased estimate with no restriction on C, while D_2 and D_3 variance estimates are only unbiased if $C = 0$. The advantage of the D_3 estimate is that it optimally weights the information about B in both x and y. The D_1 approach is reasonable for confidence interval estimation, but clearly conservative, and not in the Neyman and Pearson[3] spirit for hypothesis testing.

The fourth rule D_e, defined in Section IV, is the uniformly most powerful similar decision rule based on similar critical regions so that the significance level is always $100\alpha\%$ independent of the true background counting rate B. The exact rule D_e is more complicated to use than the normal approximation rules. Tables of critical points and randomization probabilities are included in Appendix A.

Power surfaces of D_1, D_2, D_3 and D_e are plotted in Figures 2, 3, and 4 for selected values of $\rho = t/s$ and B. These show the robustness of D_3 and compare D_1, D_2, and D_3 with the exact rule D_e. An extensive set of D_3 power surface sections is graphed in Appendix B. The limiting power of all four rules as $(C + B)t$ and Bs approach infinity is also derived in Section IV. Graphs of the limiting power surface sections are included in Appendix E.

Section V considers the answer to question Q_3 from the standpoint of confidence interval estimation. Two different confidence interval rules I_1 and I_4 for C are described and compared in detail. The approximate $100(1-\alpha)\%$ confidence forms of these rules are:

I_1. Decide that C satisfies

$$\tilde{c} - n_{\alpha/2}[x/t^2 + y/s^2]^{1/2} \leqq C \leqq \tilde{c} + n_{\alpha/2}[x/t^2 + y/s^2]^{1/2};$$

I_4. Decide that C satisfies

$$L_4(x, y, \rho, \alpha)/t \leqq C \leqq U_4(x, y, \rho, \alpha)/t.$$

For the customary confidences of 90%, 95%, and 99%, the $n_{\alpha/2}$ values are $n_{0.05} = 1.645$, $n_{0.025} = 1.960$, and $n_{0.005} = 2.576$, respectively. The L_4 and U_4 endpoint functions of I_4 are tabled in Appendix D for appropriate $x, y, \rho = t/s$ and $1-\alpha$ values. Both the rules are based on the normal approximation to the Poisson distribution, so the stated confidence coefficients are not exact.* Rule I_1 is the symmetric confidence interval formulation of rule D_1 of (2). Rule I_4 is based on a joint confidence region for B and C approach. The mathematical details of the argument constitute Appendix C.

In Section V the two rules are compared with respect to exact confidence coefficients and expected interval lengths. Tables are contained in Appendix D. Rule I_4 is seen to be robust against the departures from normality resulting from very small expected total counts. An asymptotic length formula for both rules when $(C + B)\,t$ and Bs are large is derived in Section V. A graph is included in Appendix E.

The importance of the background estimate y/s for good inference is considered in Section VI. Graphs show the effect on both power of hypothesis tests and expected length of confidence intervals of varying the background information parameter $\rho = t/s$ and the expected total background count Bt. Asymptotic results are also included.

In the first half of Section VI it is assumed that a background estimate is only used once. The remainder of the Section considers the average value characteristics of rules when the same background

---

* It is not possible to construct an exact confidence interval for C (except in the trivial instance when it is known that b = B) based on Poisson variables x and y. Alternative approaches do exist. Probably the best solution to the whole area of statistical inference based on a few counts is to change the fundamental counting procedure from one of monitoring the random number of counts in a fixed time period to one of monitoring the random time to accumulate a fixed number of counts. This latter procedure based on random time does not suffer from the discreteness limitation of the Poisson distribution. Statistical techniques for random time counting are not discussed here. A discussion of random time procedures and related sequential techniques is included in Reference[4].

estimate is used to correct a number of different sample gross counts.

II. RECOMMENDATIONS

The mathematical and numerical investigations into the three questions Q_1, Q_2, and Q_3 raised in Section I are reported in detail in later sections of this paper. The specific recommendations resulting from this work are presented below in the form of several rules of thumb. These recommendations are meant to cover a gamut of possible combinations of expected background count tB, sample to background counting times ratio ρ and level of significance $100\alpha\%$ (or confidence coefficient $1-\alpha$). As such they will have deficiencies in specific circumstances. If a priori information is available concerning tB (a range, or upper, or lower bound, for example), a specific set of ρ values is under consideration, and/or a particular α is to be used, detailed reading of later text material will undoubtedly produce a rule superior to the rules of thumb below.

The choice of a hypothesis testing rule to answer question Q_1 could best be based on ρ, sB, and the discrepancy between stated $100\alpha\%$ level of significance and true level of significance that the experimenter can tolerate. Since sB is unknown, y is substituted as a choice parameter. The rule D_3 has the smallest such discrepancy if $\rho \leqq 1$ (Table I). The discrepancy is positive and decreases as tB increases for fixed ρ. Power surface sections (Appendix B) suggest the following rule of thumb:

RT_1: For stated α equal to $\begin{Bmatrix} 0.10 \\ 0.05 \\ 0.01 \end{Bmatrix}^*$ use decision rule D_3 when $\rho \leqq 1$ and $y > \begin{Bmatrix} 5 \\ 10 \\ 15 \end{Bmatrix}$; use exact decision rule D_e when $\rho > 1$ and/or $y \leqq \begin{Bmatrix} 5 \\ 10 \\ 15 \end{Bmatrix}$.

---

* In the remainder of the statement, use the top quantity in each bracket with 0.10, the second with 0.05, etc.

In using RT_1 almost all applications of D_3 when H_o is true will have the true type I error $\bar{\alpha}$ satisfying $\begin{Bmatrix} 0.10 \\ 0.05 \\ 0.01 \end{Bmatrix} \leqq \alpha \leqq \begin{Bmatrix} 0.12 \\ 0.07 \\ 0.02 \end{Bmatrix}$. This discrepancy range would seem to be acceptable under most conditions. Tighter or looser bounds on the discrepancy are best accomplished by increasing or decreasing, respectively, the y bound for use of D_3.

The threshold C_o of question Q_2 is definable in terms of power (Section III). Thus, C_o as a function of $1 - \beta$ is the smallest expected net sample counting rate for which a particular decision rule (with specified level of significance) has power not less than $1 - \beta$. For large x and y the asymptotic power function (28) and power curves of Appendix B suggest the following rule of thumb:

RT_2: For stated $100\alpha\%$ level of significance, power $(1 - \beta) \geqq 1/2$, and expected background contribution to sample count tB the threshold C_o for decision rule D_3 satisfies

$$C_o \geqq \sqrt{tB(1 + \rho)}\ (n_\alpha + n_\beta)/t.$$

Equality is approximately satisfied for $tB \geqq 50$.

The asymptotic curves of Appendix E, Figures E.1, E.2, and E.3, give the equality form of this threshold relationship. For finite x and y the inequality is strict. The approach to the limiting form for fixed ρ is depicted in the families of power sections of Appendix B. In each figure the limiting power section is the straight line on the far left of the family of curves. The first curve to the right of the straight line corresponds to $tB = 50$.

A common practice when different counting systems are being compared is the basing of the selection of the best system on the "signal to noise ratio", or, in previously defined symbols, the ratio C/B. The present investigation suggests the alternative rule of thumb:

RT'_2: The probability of detecting a net sample counting rate C in the presence of a known average background counting rate B based on a counting period of time t increases with $(tC)/(tB)^{1/2}$.

If the expected background rate is unknown and estimated over a period of time s the detection probability increases with $(tC)/[tB(1 + \rho)]^{1/2}$.

Our rule of thumb RT_2' is quite close to the "figure of merit" C^2/B suggested by Arnold.[5] To think in terms of counting rates can be misleading if alternatives involve different counting schedules. The expected total count attributed to the sample tC and the expected total background count in the gross sample count tB are the fundamental quantities which control detection and also precision of estimating C, the subject of question Q_3.

To answer question Q_3, a confidence interval estimate of C can be constructed around $\tilde{c}$ as a "center." For $tB > 1$ and $\rho \leq 1$, both confidence intervals I_1 and I_4 have true confidence coefficients close to the stated ones. In this region I_4 is conservative in the sense that its confidence coefficient is never less than the stated one. I_1 confidence coefficients fluctuate around the stated one, being significantly low most of the time for small tB. The differences in the two confidence coefficients become neglegible for $tB > 10$. Rephrasing in terms of y gives the following rule of thumb:

RT_3: For any of the confidence coefficients 0.90, 0.95, and 0.99 and $\rho \leq 1$ use confidence interval rule I_4 if $2 + 3/2\rho \leq y \leq 6 + 12/\rho$ and use I_1 if $6 + 12/\rho < y$. Use of either I_1 or I_4 is questionable if $y < 2 + 3/2\rho$.

The conservative nature of rule I_4 is compromised somewhat by expected confidence interval lengths that are longer than those for I_1. The greatest differences are for $\rho = 1/10$ and tC and tB small. (Using the net count form of the intervals, the greatest difference is about 4 with $1 - \alpha = 0.99$, about 2.5 with $1 - \alpha = 0.95$, and about 1.3 with $1 - \alpha = 0.90$). As either $1/\rho$, tC, or tB increases the differences decrease. The rule of thumb RT_3 is the result of the author's preference for a rigid lower bound on the confidence over a shorter expected length.

The theoretical and numerical results of Sections IV and V were reevaluated in Section VI to determine the importance of the background estimate for high power in hypothesis tests and for short confidence intervals. The results are summarized in the following rule of thumb:

RT_4: If a choice is available as to the allocation of a fixed total counting time between sample and background (where the background is not to be used for other samples), split the counting time equally between sample and background ($\rho = 1$). If a background count is needed to match a sample counted for a fixed time t, the background counting time s should always satisfy $s \geqq t$; only in extreme situations should $s > 3t$.

In the last part of Section VI it is shown that the average value properties of hypothesis tests and confidence intervals do not change when backgrounds are used to correct several samples. However, the variance of the fraction of correct decisions and the average confidence interval length both increase with the number of times the background is used, because answers tend to be blocked on background. Numerical investigation of the asymptotic variance for hypothesis tests suggests the following rule of thumb for combating this effect by increasing the precision of the background estimate:

RT_5: Suppose each background estimate, based on a count over time s, is to be used to correct M independent samples, each based on count over time t. ρ and M should be inversely related in agreement with the following:

For $M \leqq 10$, have $\rho \leqq 1$;
For $M = 25$, have $\rho \approx 1/4$;
For $M = 75$, have $\rho \approx 1/10$.

As indicated in Section I the investigations reported in the paper are concerned with the region of appreciable background correction, defined as $0 \leqq k = (tC)/(tB)^{1/2} < 15$. The above rules of thumb only apply in this region. For large k background is a nominal problem. A reasonable rule of thumb which insures sufficient background information without being overly conservative is the familiar rule which minimizes the variance of $\widetilde{c}$:

RT_6: For $k \geqq 15$ and either hypothesis testing or confidence interval applications select

$$\rho \approx (B + C)^{1/2}/B^{1/2}.$$

III. STATISTICAL CONCEPTS

To insure that the reader is familiar with the statistical concepts needed to discuss the questions raised in the introduction, the following description of hypothesis testing and confidence interval estimation is included. For a more detailed treatise of these concepts the reader is referred to any of a number of standard statistical tests; e.g., [6, 7, 8]. The description below is purposely couched in terms of the specific application dealt with in the paper.

Let x and y be independent Poisson chance variables with mean values $(B + C)t$ and Bs, respectively. The sample space S for x and y, the set of all possible experimental values that can be taken on by x and y, is the infinite lattice of points (x, y) with x and y non-negative integers. The joint probability density function (p.d.f.) for x and y defined on the points of S has the form

$$\{\exp[-(B + C)t - Bs]\} [(B + C)t]^x (Bs)^y / x!\, y!\, . \qquad (3)$$

Let Ω be the closed first quadrant of two-dimensional space. With t and s fixed by the experimental situation there is a one-one correspondence between the possible p.d.f.'s for x and y and the points (B, C) of Ω. The set Ω is called the "parameter space." Let ω be the subset of Ω consisting of all points of the form $(B, 0)$.

The null hypothesis H_o of Section I now states that the true p.d.f. in (3) for x and y corresponds to a point of ω, or simply "is a point of ω." The alternative hypothesis H_C states that the true p.d.f. is a point of $\Omega - \omega$. A decision rule for answering question Q_1 is a mapping of the sample space S on to the parameter space Ω. S is divided into two sets, say S_o and S_C. If the experimental point (x, y) falls in S_o decision H_o or ω is made. However, if it falls in S_C decision H_C or $\Omega - \omega$ is made.

In the Neyman and Pearson theory of hypothesis testing[3] the goodness of a decision rule is measured by the probability with which it makes wrong decisions. Two different wrong decisions are possible depending upon the true value of C. However, in any given application it is only possible to make one of these errors. If H_o is true, the only possible error is to decide H_C is true. This error is called "alpha" or "type I." The probability of an alpha error occurring is designated by a lower case α. Symbolically,

$$\begin{aligned} \alpha(B,t,s) &= \Pr\ \{(x,y)\epsilon S_C | C = 0, B, t, s\} \\ &= \Pr\ \{H_C | \omega\}. \end{aligned} \tag{4}$$

Here, the mathematical notation $\Pr\{A|B\}$ stands for the probability that the event A will occur given that the state of affairs is B. The familiar term "level of significance" refers to α expressed in percent; i.e., the level of significance is $100\alpha\%$. In general, the alpha error for a decision rule will depend upon B, t, and s. This dependence is indicated in (4) above. If H_C is true, an error is committed only when decision H_o is made. Such an error is called "beta" or "type II." The probability of a beta error occurring is analogously designated by a lower case β. As above

$$\begin{aligned} \beta(C,B,t,s) &= \Pr\ \{(x,y)\epsilon S_o | C, B, t, s\} \\ &= \Pr\ \{H_o | \Omega - \omega\}. \end{aligned} \tag{5}$$

In addition to depending on B, t, and s, β is a function of C. For a reasonable decision rule, β would be a decreasing function of C; i.e., the larger that C is, the less likely that a beta error would be committed. Also, one would expect that as $C \to +\infty$, $\beta(C, B, t, s) \to 0$ for every fixed set (B, t, s). The compliment of β, $1 - \beta$, is called the "power" of the decision rule. Clearly,

$$1 - \beta(0, B, t, s) = \alpha(B, t, s). \tag{6}$$

The complete picture of the probabilities of errors, both types I and II, is depicted by a three-dimensional plot of the power surface over Ω. A typical power surface appears in Figure 1.

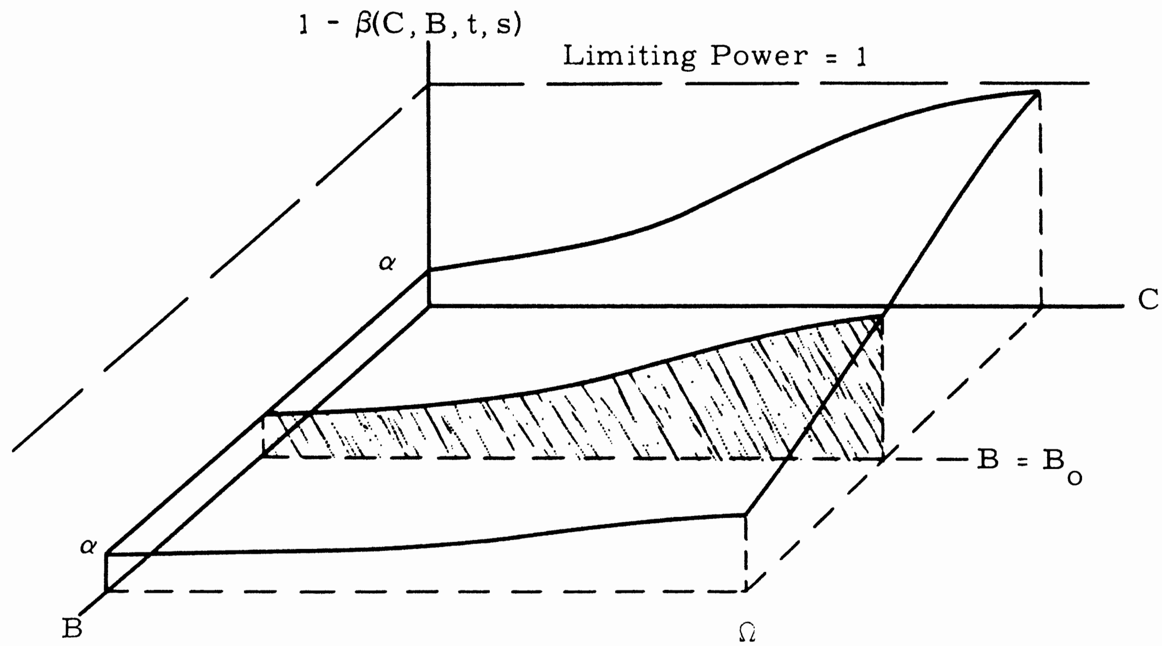

FIGURE 1

Typical Power Surface for a Decision Rule Test of H_o Against H_C

For the moment suppose that the background counting rate in our problem is known to be $B = B_o$; i.e., the parameter space is one-dimensional consisting of all points (B_o, C). The structures of the two hypotheses H_o and H_C are quite different. H_o completely specifies the distribution of the chance variables x and y. As such it is called a "simple" hypotheses. H_C merely states that the distribution of x is one of an infinite family with C positive. Or in other words, H_C is the union of an infinite number of simple hypotheses. The distinction is made by calling H_C a "composite" hypothesis. The Neyman and Pearson selection criterion for a best decision rule is that both alpha and beta errors be nominal in the following sense.

First, a tolerable alpha error α is specified, say $\alpha = 0.05$, a 5% level of significance. Usually an infinite family of decision rules will have an alpha error not greater than 0.05. Second, from among this family of rules one is selected which makes β small. Since β is a function of C, one particular value of C, say C_1, is considered. Any rule which minimizes $\beta(C_1, B, t, s)$ over the entire family is called a "most powerful" decision rule for C_1 at level α, since minimizing β is equivalent to maximizing power at C_1. Usually, the most powerful rule changes with C_1. The cases when the Neyman and Pearson approach is really fruitful are those for which the most powerful rule is independent of C_1. Such a rule is called "uniformly most powerful." Refering to Figure 1 the depicted decision rule would be uniformly most powerful for $B = B_o$ if above the line $B = B_o$ in Ω the surface were the envelope of all possible decision rules with the same type I error. If one accepts the thesis that minimizing alpha and beta errors is the correct way to select a decision rule and, if the rule most suffice for a range of C values, a uniformly most powerful rule would seem to be the best that one could hope for.

Our problem is complicated because the parameter space Ω is two dimensional. The uniformly most powerful rule for $B = B_o$ is a function of B_o. So, with B unknown, a uniformly most powerful rule does not exist. In general, the level of significance of any decision rule changes with B. For our problem there are rules which have constant level of significance independent of B. (The power surface in Figure 1 is for such a rule.) These rules are called "similar rules" or "rules with Neyman structure" after Neyman who first discussed their existence(9). For each type I error α there is a unique uniformly most powerful similar rule for our problem.

In Section IV four different decisions rules D_1, D_2, D_3 and D_e are described and their power surfaces compared. The first three rules, D_1, D_2, and D_3, are based on the normal approximation to the Poisson distribution. Nominal levels of significance, 1%, 5%, and 10%, are set for these rules using normal distribution theory. The actual levels fluctuate around the nominal ones as B, t, and s vary. The selection of a best rule from among the set of three entails not only a comparison of the power surfaces but also a comparison of the degree of fluctuation of the actual significance levels about the nominal ones.

The attraction of rules based on a normal approximation is that of simplicity. Tables are not needed for application. As a reference the exact significance level rule D_e is evaluated also. Since D_e is the uniformly most powerful similar rule for all alternative hypotheses H_C $(C > 0)$, an optimum standard is provided against which to measure the power characteristics of the simple rules D_1, D_2, and D_3.

In confidence interval estimation a random interval $I(x, y, t, s)$ is constructed with the property that the probability that $I(x, y, s, t)$ covers the true net sample counting rate C is a fixed fraction $1 - \alpha$, independent of the actual value of C. It is important to realize that it is the interval, not the unknown count rate C, which is random. After an experimental observation (x, y, t, s) has been taken and a particular interval I(x, y, t, s) has been constructed there is no probability involved. The parameter C, though unknown, is an absolute quantity, not subject to random fluctuations. Either it has been covered by I or it has not. The term "confidence" describes the goodness of the covering I in an average frequency sense. In a large number of independent applications of the random interval technique approximately $100(1 - \alpha)\%$ of the intervals will actually cover the true value of C. While a definite statement concerning coverage on a particular application is impossible, it can be said that on the average coverage occurs about $100(1 - \alpha)\%$ of the time. This average frequency property of the coverage interval is described on a single application by the statement that one is $100(1 - \alpha)\%$ confident that the parameter C lies in the interval, or that I is a $100(1 - \alpha)\%$ confidence interval for the true value of C. Here, $(1 - \alpha)$ is termed the "confidence coefficient."

Confidence interval decision rules and hypothesis testing decision rules are closely related. In fact, every confidence interval rule I defines a hypothesis testing rule D. The definition is set up in the following manner.

> For every experimental data point (x, y, t, s), D makes decision "H_o is true" if and only if I covers 0.

Thus, type I error and confidence coefficient are numerically complimentary fractions for I and the corresponding D.

In general, there will be an infinite number of confidence interval rules with the same confidence coefficient. The goodness of a rule can be measured by the average length of the interval. This average length will usually increase with C. A best rule for fixed $C = C_1$ would be one which gives the shortest average length. It would be uniformly best if it gave the shortest average length for all $C \geqq 0$.

As for hypothesis testing rules the fact that the background counting rate B is unknown complicates the evaluation of confidence interval rules. Two rules I_1 and I_4 are discussed in Section V. Both are based on the normal approximation to the Poisson distribution. Stated confidence coefficients of 0.90, 0.95, and 0.99 define the rules. Evaluation consists of investigating the true confidence and average length as functions of B, C, t and s.

It is not possible to construct an exact confidence interval rule for C when B is unknown so there is no optimum standard with which to compare I_1 and I_4. An exact rule for the ratio C/B does exist (see Chapman[10]) which is the confidence interval rule that defines the hypothesis testing rule D_e discussed in Section IV.

IV. HYPOTHESIS TESTS

Let x and y be independent Poisson chance variables with mean values t(B + C) and sB, respectively. With t and s known there is a one-one correspondence between the possible p.d.f.'s for x and y and the parameter space Ω consisting of all points (B, C) in the closure of the first quadrant of the two-dimensional Cartesian plane. Let ω be the subset of Ω consisting of all points of the form (B, 0). The problem is to construct a decision rule, based only on the knowledge of x, y, t, and s to test the null hypothesis,

$$H_o : C = 0, \tag{7}$$

(i.e., that the true parameter point lies in ω) against the alternative hypothesis,

$$H_C : C > 0, \tag{8}$$

(i.e., that the true parameter point lies in $\Omega - \omega$).

Normal Approximation Decision Rules

The simplest approach to the solution of the above problem is to assume that x and y are normally distributed, construct a standard normal zero-one statistic, substitute estimates for parameters where needed, and test whether the resulting statistic is significantly larger than zero assuming a normal zero-one distribution. This approach gives a decision rule of the form

$$D: \text{ Decide } H_C \text{ is true if } x - \rho y > n_\alpha \tilde{\sigma}; \qquad (9)$$
$$\text{Otherwise, decide } H_o \text{ is true.}$$

Here, $\rho = t/s$, the ratio of the means when H_o is true. The quantity n_α is the $100(1 - \alpha)$ percentile of the normal distribution with mean zero and variance one; i.e., n_α satisfies $F(n_\alpha) = 1 - \alpha$, where F is the cumulative normal distribution function defined as

$$F(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-a^2/2} \, da. \qquad (10)$$

The quantity $\tilde{\sigma}$ is an estimate of the standard deviation σ of $x - \rho y$.

It is well known (e.g., Feller[11], p. 230) that the normality assumption is valid in the limit as $t(B + C) \to +\infty$ and $sB \to +\infty$. If both means $t(B + C)$ and sB are sufficiently large and if $\tilde{\sigma}$ is a good estimate of σ, the normal approximation approach should be almost exact in the sense that the actual type I error will be about equal to the stated one α. In addition, the test should have uniformly high power since the corresponding normal distribution test is uniformly most powerful. In the sequel three normal approximation tests are considered, each based on a different $\tilde{\sigma}$ estimate. Their power surfaces are investigated to determine the goodness of the type I error approximation for small mean values and to determine the regions of the parameter space Ω where one test is powerwise superior to the others.

The variance σ^2 of $x - \rho y$ is

$$\begin{aligned} \sigma^2 &= t(B + C) + \rho^2(sB) = (tC) + \rho(1 + \rho)(sB) \\ &= (tC) + (1 + \rho)(tB). \end{aligned} \qquad (11)$$

From the first line of (11) the most obvious unbiased estimate of σ^2 is

$$\tilde{\sigma}_1^2 = x + \rho^2 y. \qquad (12)$$

The estimate $\tilde{\sigma}_1$ gives the decision rule

D_1. Decide H_C is true if $x - \rho y > n_\alpha[x + \rho^2 y]^{1/2}$; (13)

Otherwise, decide H_o is true.

Statement (13) will be called a "net count" formulation of the decision rule D_1 because in the physical application $x - \rho y$ is the sample net count estimate. Division of the inequality in D_1 by t gives the "net count rate" form of the rule. The net count rate forms of all the rules are stated in (2) of the introduction.

Use of the variance estimate $\tilde{\sigma}_1^2$ in a decision rule with null hypothesis H_o is not in agreement with the classical Neyman and Pearson theory of hypothesis testing.[3] Strictly speaking, the <u>null</u> hypothesis variance of $x - \rho y$ belongs in the decision rule. In lieu of this, an estimate should be used which is not seriously affected by the introduction of a positive sample ($C > 0$) into the null background-only distribution. For $\rho \leq 1$, the case of practical interest, D_1 places most of the weight in the variance estimate on x which causes C to affect the decision rule in the worst possible manner. Accordingly, the power of D_1 suffers.

From (11) the simplest estimate of σ^2 when $C = 0$ is

$$\tilde{\sigma}_2^2 = \rho(1 + \rho)y, \qquad (14)$$

which gives the decision rule

D_2. Decide H_C is true if $x - \rho y > n_\alpha[\rho(1 + \rho)y]^{1/2}$; (15)

Otherwise, decide H_o is true.

When C = 0, x by itself possesses information about B; in fact, the mean of x is just tB. The proper linear combination of x and y will give a better estimate of σ than either $\tilde{\sigma}_1$ or $\tilde{\sigma}_2$. Now, x and ρy are independent unbiased estimates of tB with variances (tB) and ρ (tB), respectively. Thus,

$$\left[\frac{x}{tB} + \frac{\rho y}{\rho(tB)}\right] \Big/ \left[\frac{1}{tB} + \frac{1}{\rho(tB)}\right] = \frac{\rho\,(x+y)}{1+\rho}, \tag{16}$$

is the minimum variance, unbiased, linear estimate of tB based on x and y. The result (16) gives

$$\tilde{\sigma}_3^2 = \rho(x+y) \tag{17}$$

as the minimum variance, unbiased, linear estimate of σ^2 based on x and y when C = 0. The net count form of the corresponding decision rule is

D_3. Decide H_C is true if $x - \rho y > n_\alpha[\rho(x+y)]^{1/2}$; (18)

Otherwise, decide H_o is true.

An interesting property of D_3 is its symmetric treatment of sample and background information - a property not shared by either D_1 or D_2. Of course, this symmetry is a direct consequence of the optimum weighting of the two counts as illustrated in (16). If the background count is taken over a very long time period so that y and s dominate the variance calculation (i.e., $y \xrightarrow{Pr} +\infty$, $s \rightarrow +\infty$, $y/s \xrightarrow{Pr} B$), the rule D_3 reduces to a test of whether the sample counting rate x/t is a normal variable with mean B and variance B/t. Or, in other words, it is a test of whether x/t is too large to be an observation from the known background distribution. On the other hand, if the sample count dominates (i.e., $x \xrightarrow{Pr} +\infty$, $t \rightarrow +\infty$, $x/t \xrightarrow{Pr} B + C$), D_3 reduces to a test of whether the background counting rate y/s is too small to be an observation from the known sample total count distribution.

For a given value of ρ, tB and α the exact significance levels of the decision rule D_i $(i = 1, 2, 3)$, say $\alpha_i(\rho, tB, \alpha)$, is

$$\alpha_i(\rho, tB, \alpha) = \Pr \{(x, y) \in S_i \mid \rho, tB, \alpha, C = 0\} , \tag{19}$$
$$= [\exp -(1+\rho)tB] \sum_{(x,y)\in S_i} (tB)^x (\rho tB)^y / x!\,y! \ .$$

The set S_i is the subset of the sample space where the points (x, y) satisfy $x - \rho y > n_\alpha \tilde{\sigma}_i$. The calculation (19) was done on an IBM 7090 for a representative set of ρ, tB and α values. Some of these results for the three decision rules D_1, D_2, and D_3 and stated significance levels of 1%, 5%, and 10% are reported in Table I. For each significance level the $\rho = 0$ entries are the limiting probabilities when the background counting rate B is known exactly. (This is mathematically equivalent to a background estimate based on an infinite counting period.) In the limit D_2 and D_3 are identical. D_1 neglects the variance information in B, using $x^{1/2}$ as a variance estimate.

The comparison of the significance levels of D_1, D_2, and D_3 in Table I clearly shows the robustness of D_3. The pattern of the exact levels is similar for all three levels, $\alpha = 0.01$, 0.05, and 0.10. Dropping the argument of the α_i functions defined in (19), the fundamental characteristics of the patterns seem to be:

1. For $0 < \rho \leqq 1/2$, α_3 is uniformly closer to α than either α_1 or α_2. The ordering of the errors is $\alpha_1 < \alpha < \alpha_3 < \alpha_2$. The biases in α_1 and α_2 are extreme for small tB and $\alpha = 0.01$; both biases decrease as tB increases, while α_1 bias decreases and α_2 increases as α and/or ρ increases. Only for isolated instances with tB = 1, 2 is α_3 badly biased.

TABLE I

EXACT SIGNIFICANCE LEVELS FOR DECISION RULES D_1, D_2, AND D_3

Stated Level of Significance: 1%(n_α = 2.326)

Expected Total Background Contribution - (tB)

ρ = t/s		1	2	3	5	10	15	20	30	40	50
0	D_1	0.000	0.000	0.000	0.001	0.002	0.002	0.003	0.004	0.004	0.004
	D_2	0.018	0.017	0.012	9.014	0.014	0.019	0.013	0.015	0.014	0.012
	D_3	0.018	0.017	0.012	0.014	0.014	0.019	0.013	0.015	0.014	0.012
1/10	D_1	0.000	0.000	0.000	0.001	0.002	0.003	0.003	0.004	0.005	0.005
	D_2	0.042	0.029	0.025	0.022	0.018	0.017	0.016	0.014	0.014	0.013
	D_3	0.024	0.022	0.020	0.017	0.015	0.014	0.014	0.013	0.013	0.012
1/5	D_1	0.000	0.000	0.001	0.001	0.003	0.003	0.004	0.005	0.005	0.006
	D_2	0.041	0.038	0.031	0.023	0.020	0.018	0.017	0.016	0.015	0.014
	D_3	0.019	0.017	0.018	0.017	0.014	0.013	0.013	0.012	0.012	0.012
1/3	D_1	0.000	0.001	0.001	0.002	0.003	0.004	0.005	0.006	0.006	0.006
	D_2	0.094	0.051	0.036	0.029	0.021	0.020	0.018	0.017	0.016	0.015
	D_3	0.021	0.019	0.016	0.014	0.012	0.013	0.012	0.012	0.012	0.012
1/2	D_1	0.000	0.001	0.002	0.003	0.005	0.005	0.006	0.007	0.007	0.007
	D_2	0.113	0.072	0.052	0.037	0.026	0.022	0.020	0.018	0.017	0.016
	D_3	0.012	0.013	0.013	0.013	0.012	0.012	0.011	0.011	0.011	0.011
1	D_1	0.000	0.003	0.006	0.009	0.009	0.010	0.010	0.010	0.010	0.010
	D_2	0.234	0.133	0.083	0.058	0.036	0.031	0.027	0.022	0.020	0.019
	D_3	0.000	0.003	0.006	0.009	0.009	0.010	0.010	0.010	0.010	0.010
2	D_1	0.000	0.006	0.019	0.033	0.022	0.021	0.019	0.016	0.016	0.015
	D_2	0.383	0.319	0.216	0.109	0.058	0.046	0.039	0.030	0.027	0.025
	D_3	0.000	0.000	0.000	0.001	0.005	0.006	0.006	0.007	0.008	0.008

TABLE I (Continued)

EXACT SIGNIFICANCE LEVELS FOR DECISION RULES D_1, D_2, AND D_3

Stated Level of Significance: 5%(n_α = 1.645)

Expected Total Background Contribution - (tB)

ρ = t/s		1	2	3	5	10	15	20	30	40	50
0	D_1	0.004	0.005	0.012	0.014	0.027	0.033	0.034	0.032	0.039	0.032
	D_2	0.080	0.053	0.084	0.068	0.049	0.053	0.052	0.046	0.053	0.056
	D_3	0.080	0.053	0.084	0.068	0.049	0.053	0.052	0.046	0.053	0.056
1/10	D_1	0.005	0.011	0.015	0.020	0.027	0.030	0.032	0.035	0.037	0.038
	D_2	0.086	0.076	0.071	0.066	0.061	0.060	0.059	0.057	0.056	0.055
	D_3	0.067	0.064	0.063	0.060	0.058	0.056	0.055	0.054	0.054	0.054
1/5	D_1	0.007	0.015	0.018	0.023	0.030	0.034	0.034	0.037	0.038	0.040
	D_2	0.107	0.086	0.077	0.069	0.065	0.065	0.060	0.958	0.057	0.057
	D_3	0.069	0.062	0.059	0.057	0.054	0.055	0.055	0.054	0.053	0.053
1/3	D_1	0.009	0.016	0.023	0.029	0.032	0.035	0.037	0.040	0.041	0.042
	D_2	0.111	0.095	0.084	0.078	0.069	0.064	0.062	0.060	0.059	0.058
	D_3	0.094	0.059	0.056	0.057	0.055	0.054	0.054	0.053	0.052	0.052
1/2	D_1	0.017	0.028	0.030	0.031	0.037	0.039	0.040	0.042	0.043	0.044
	D_2	0.166	0.122	0.107	0.083	0.071	0.069	0.066	0.063	0.061	0.060
	D_3	0.063	0.070	0.062	0.053	0.052	0.053	0.052	0.053	0.052	0.052
1	D_1	0.030	0.050	0.050	0.052	0.049	0.050	0.050	0.050	0.051	0.050
	D_2	0.240	0.160	0.122	0.098	0.087	0.079	0.073	0.070	0.068	0.065
	D_3	0.030	0.050	0.050	0.052	0.049	0.050	0.050	0.050	0.051	0.050
2	D_1	0.049	0.110	0.133	0.101	0.073	0.069	0.065	0.063	0.061	0.060
	D_2	0.383	0.320	0.224	0.139	0.116	0.095	0.090	0.081	0.076	0.072
	D_3	0.000	0.006	0.019	0.038	0.040	0.042	0.046	0.047	0.046	0.047

TABLE I (Continued)

EXACT SIGNIFICANCE LEVELS FOR DECISION RULES D_1, D_2, AND D_3

Stated Level of Significance: 10%(n_α = 1.282)

Expected Total Background Contribution - (tB)

ρ = t/s		1	2	3	5	10	15	20	30	40	50
0	D_1	0.019	0.053	0.034	0.068	0.083	0.083	0.078	0.089	0.092	0.092
	D_2	0.080	0.143	0.084	0.133	0.083	0.125	0.112	0.089	0.092	0.092
	D_3	0.080	0.143	0.084	0.133	0.083	0.125	0.112	0.089	0.092	0.092
1/10	D_1	0.031	0.042	0.050	0.060	0.071	0.075	0.078	0.082	0.084	0.086
	D_2	0.117	0.120	0.113	0.111	0.108	0.107	0.106	0.105	0.105	0.104
	D_3	0.114	0.113	0.113	0.107	0.106	0.105	0.103	0.103	0.102	0.102
1/5	D_1	0.033	0.050	0.059	0.064	0.074	0.079	0.082	0.084	0.087	0.087
	D_2	0.129	0.127	0.128	0.118	0.116	0.109	0.109	0.107	0.107	0.106
	D_3	0.116	0.112	0.106	0.106	0.104	0.104	0.103	0.102	0.102	0.102
1/3	D_1	0.048	0.055	0.059	0.074	0.080	0.081	0.085	0.088	0.090	0.090
	D_2	0.154	0.127	0.118	0.124	0.115	0.115	0.111	0.110	0.108	0.108
	D_3	0.112	0.115	0.115	0.105	0.104	0.101	0.103	0.102	0.101	0.101
1/2	D_1	0.063	0.072	0.075	0.079	0.085	0.089	0.090	0.091	0.092	0.093
	D_2	0.183	0.155	0.146	0.134	0.121	0.117	0.114	0.111	0.111	0.110
	D_3	0.116	0.102	0.108	0.108	0.104	0.101	0.102	0.101	0.101	0.101
1	D_1	0.104	0.124	0.115	0.101	0.101	0.101	0.101	0.101	0.100	0.099
	D_2	0.263	0.220	0.185	0.148	0.133	0.125	0.123	0.119	0.116	0.114
	D_3	0.104	0.124	0.115	0.101	0.101	0.101	0.101	0.101	0.100	0.099
2	D_1	0.160	0.220	0.190	0.136	0.122	0.122	0.119	0.114	0.111	0.111
	D_2	0.384	0.324	0.241	0.179	0.155	0.141	0.140	0.134	0.127	0.123
	D_3	0.012	0.053	0.083	0.091	0.093	0.097	0.097	0.097	0.098	0.099

2. For $\rho = 1$, $\alpha_3(\alpha_1)$ is uniformly closer to α than α_2. With $\alpha = 0.01$ or 0.05 the ordering is $\alpha_3(\alpha_1) < \alpha < \alpha_2$, while for $\alpha = 0.10$ it is $\alpha < \alpha_3(\alpha_1) < \alpha_2$. The bias in α_2 is increased in the same monotone pattern stated above. Only for tB = 1, 2 and $\alpha = 0.01$ is $\alpha_3(\alpha_1)$ badly biased.

3. For $\rho = 2$, the picture is not so orderly. The bias in α_1 is decreased over that for $\rho = 1$, and that in α_2 increased. α_3 has an extreme negative bias for small tB.

From a practical standpoint, since tB is not known exactly, background and counting time ratio ranges must be determined over which the exact significance level does not deviate excessively from the stated one. Of course, what constitutes an excessive deviation must be left to the judgment of the user. A reasonable scale of acceptability would seem to be that for $\alpha = 0.01$ exact levels in the range 0.005 to 0.02 are acceptable, for $\alpha = 0.05$ exact levels in the range 0.035 to 0.065 are acceptable and for $\alpha = 0.10$ exact levels in the range 0.08 to 0.12 are acceptable. Based on this scale of acceptability and on the fact that negative bias in the significance level is coupled with a general depression in the power function (see Figures 2, 3, 4), D_3 is certainly superior to D_1 and D_2.

Uniformly Most Powerful Similar Decision Rule

When x and y are independent Poisson chance variables with mean values $t(B + C)$ and sB, respectively, the conditional distribution of x given the sum $x + y$ is binomial on $x + y$ trials and success probability $\rho(R + 1)/[\rho(R + 1) + 1]$, where $R = C/B$ and $\rho = t/s$ (e.g., Chapman[10]). Since the parameters of this binomial distribution depend only on the sample information x, y, s, and t and the true counting rate ratio C/B, an exact decision rule for the null hypthesis H_o can be constructed from tables of the binomial distribution.

The mathematical details are contained in Reference(10) and will not be discussed here. This decision rule D_e has the form

$$
\begin{aligned}
D_e.\quad & \text{Decide } H_C \text{ is true if } x > X_\alpha(x+y, \rho); \\
& \text{Decide } H_C \text{ is true with probability} \\
& P_\alpha(x+y, \rho) \text{ if } x = X_\alpha(x+y, \rho); \\
& \text{Otherwise, decide } H_o \text{ is true.}
\end{aligned}
\tag{20}
$$

The critical point $X_\alpha(x+y, \rho)$ and the probability $P_\alpha(x+y, \rho)$ are fixed numbers which are determined by α, $x + y$, and ρ.

D_e is an example of a randomized decision rule. When $x = X_\alpha$ the decision is based on a random experiment. Figuratively, a coin is tossed with the probability of heads equal to P_α. If the coin lands heads, decision H_C is made; if it lands tails decision H_o is made. Randomization is an academic device to assure that the significance level of the rule is exactly equal to $100\alpha\%$. Since the binomial distribution is discrete, in general, the best that can be done is to select a critical point X_α which brackets the stated type I error α; i.e., X_α satisfies

$$\Pr\{x > X_\alpha \mid H_o\} < \alpha < \Pr\{x \geqq X_\alpha \mid H_o\}. \tag{21}$$

When (21) is the case, the rule D_e with

$$P_\alpha(x+y, \rho) = [\alpha - \Pr\{x > X_\alpha \mid H_o\}] / \Pr\{x = X_\alpha \mid H_o\} \tag{22}$$

has a type I error equal to α.

The conditional distribution of x given $x + y$ and $C = 0$ does not depend upon B. So, the rule D_e has a type I error identical to α on ω. The conditional binomial test on each diagonal $x + y$ = constant is the uniformly most powerful test of the null hypothesis that the success probability is $\rho/(1+\rho)$ against the one sided alternative that it is larger than $\rho/(1 + \rho)$. Therefore, the rule D_e is the uniformly most powerful similar decision rule for testing the null hypothesis H_o against alternative H_C ($C > 0$). Appendix A lists the critical points X_α and the randomization probabilities P_α for $\alpha = 0.01$, 0.05, and 0.10, and a representative range of $x + y$ and ρ values.

Since the biomial distribution with success probability q and trial parameter m is approximated by a normal distribution with mean mq and variance mq(1 - q) for m large, the exact rule D_e is asymptotically equivalent to the rule D_e'.

D_e'. Decide H_C is true

$$\text{if } x > (x + y)\frac{\rho}{\rho + 1} + n_\alpha\left[\frac{\rho}{\rho + 1}\left(1 - \frac{\rho}{\rho + 1}\right)(x + y)\right]^{1/2}; \tag{23}$$

Otherwise, decide H_o is true.

Here, n_α is the 100(1 - α) percentile of the unit normal distribution. Rewriting the inequality of (23) shows that D_e' is identical to D_3 of (18) Thus, D_3 is the asymptotic form of the uniformly most powerful similar rule D_e for t(B + C) and sB large.

Comparison of Power Surfaces

The exact power surface for the Rule $D_i(i = 1, 2, 3, e)$, say $1 - \beta_i(\rho, tB, C, \alpha)$, is

$$\begin{aligned} 1 - \beta_i(\rho, tB, C, \alpha) &= \Pr\{(x,y)\epsilon S_i \mid \rho, tB, C, \alpha\} \\ &= \{\exp[- t(B + C) - \rho tB]\} \times \\ &\sum_{\substack{x = 0 \\ y = 0}}^{\infty} M_i(x, y \mid \rho, \alpha)[t(B + C)]^x(\rho tB)^y/x!\,y!. \end{aligned} \tag{24}$$

For rules D_1, D_2, D_3, the function $M_i(x, y\ \rho, \alpha)$ satisfies

$$M_i(x, y\ \rho, \alpha) = 1 \text{ if } (x - \rho y) > n_\alpha \tilde{\sigma}_i; \qquad (25)$$
$$= 0 \text{ Otherwise.}$$

Because of the randomization nature of D_e, M_e satisfies

$$M_e(x, y\ \rho, \alpha) = 1 \text{ if } x > X_\alpha(x + y,\ \rho);$$
$$= P_\alpha(x + y, \rho) \text{ if } x = X_\alpha(x + y, \rho); \qquad (26)$$
$$= 0 \text{ Otherwise.}$$

Table I gives selected values of $1 - \beta_i(\rho, tB, 0, \alpha)$ for $i = 1, 2, 3$. By construction $1 - \beta_e(\rho, tB, 0, \alpha) \equiv \alpha$. These power surfaces were also evaluated numerically at selected points of $\Omega - \omega$ using an IBM 7090 to calculate the series in (24). Specifically, for each of the four decision rules 18 power surfaces, the (ρ, α) pairs for $\rho = 1/10$, $1/5$, $1/3$, $1/2$, 1, and 2 and $\alpha = 0.01$, 0.05, and 0.10, were investigated numerically. For each surface enough points were calculated to construct ten background constant graphical sections. Figures 2, 3, and 4 compare similar power surface sections of the four decision rules for a representative set of (ρ, α, tB) triples. The entire set of 180 power surface sections for rule D_3 appears in Appendix B.

For direct comparison with the asymptotic portion of the power surface with $t(B + C)$ and sB large the abscissa scales for all power surface sections are expressed in standardized total net sample count units k where

$$k = (tC)/(tB)^{1/2}. \qquad (27)$$

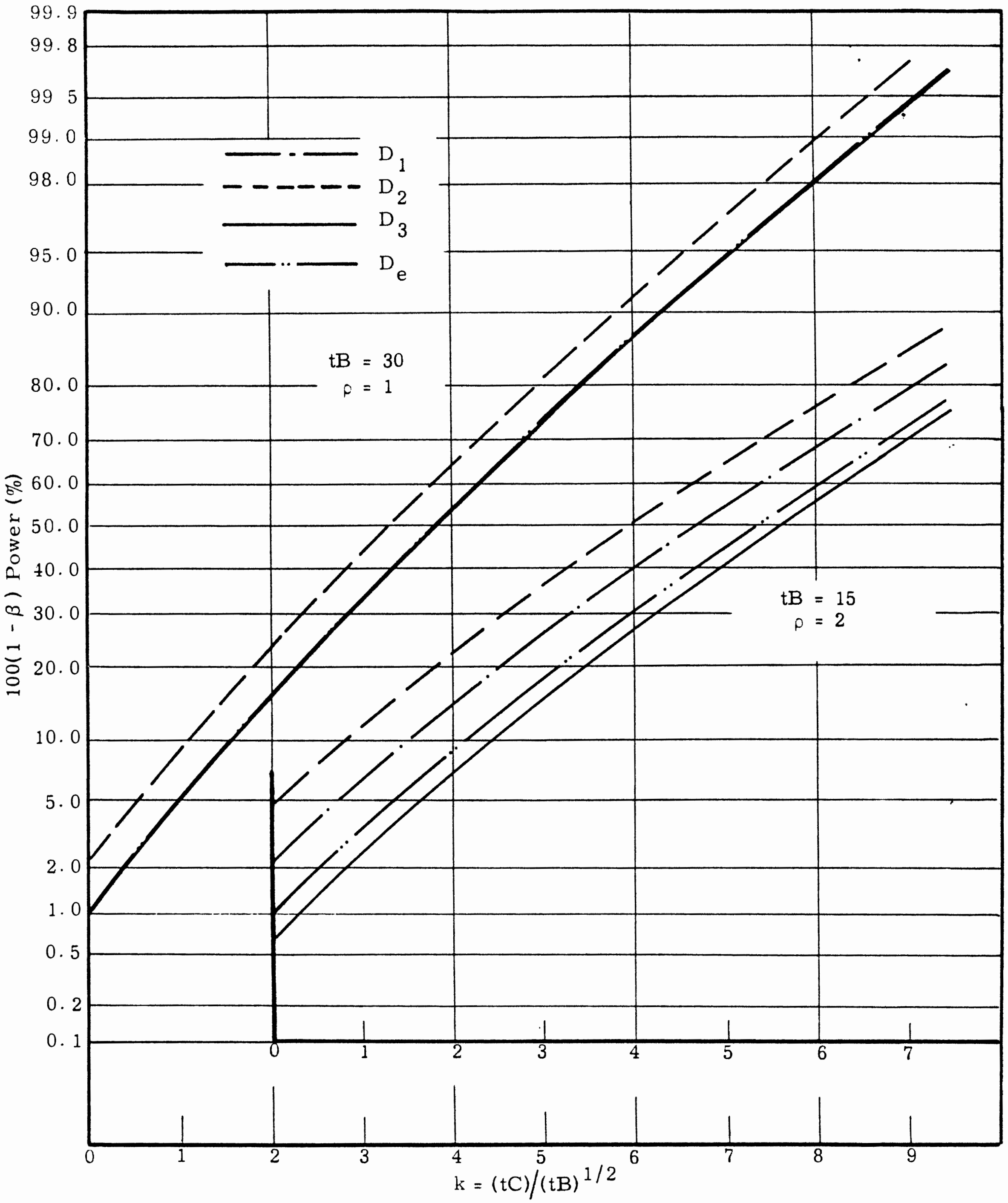

FIGURE 2

Power Surface Sections for Stated 1% Level of Significance Decision Rules D_1, D_2, D_3, and D_e for Selected tB and ρ

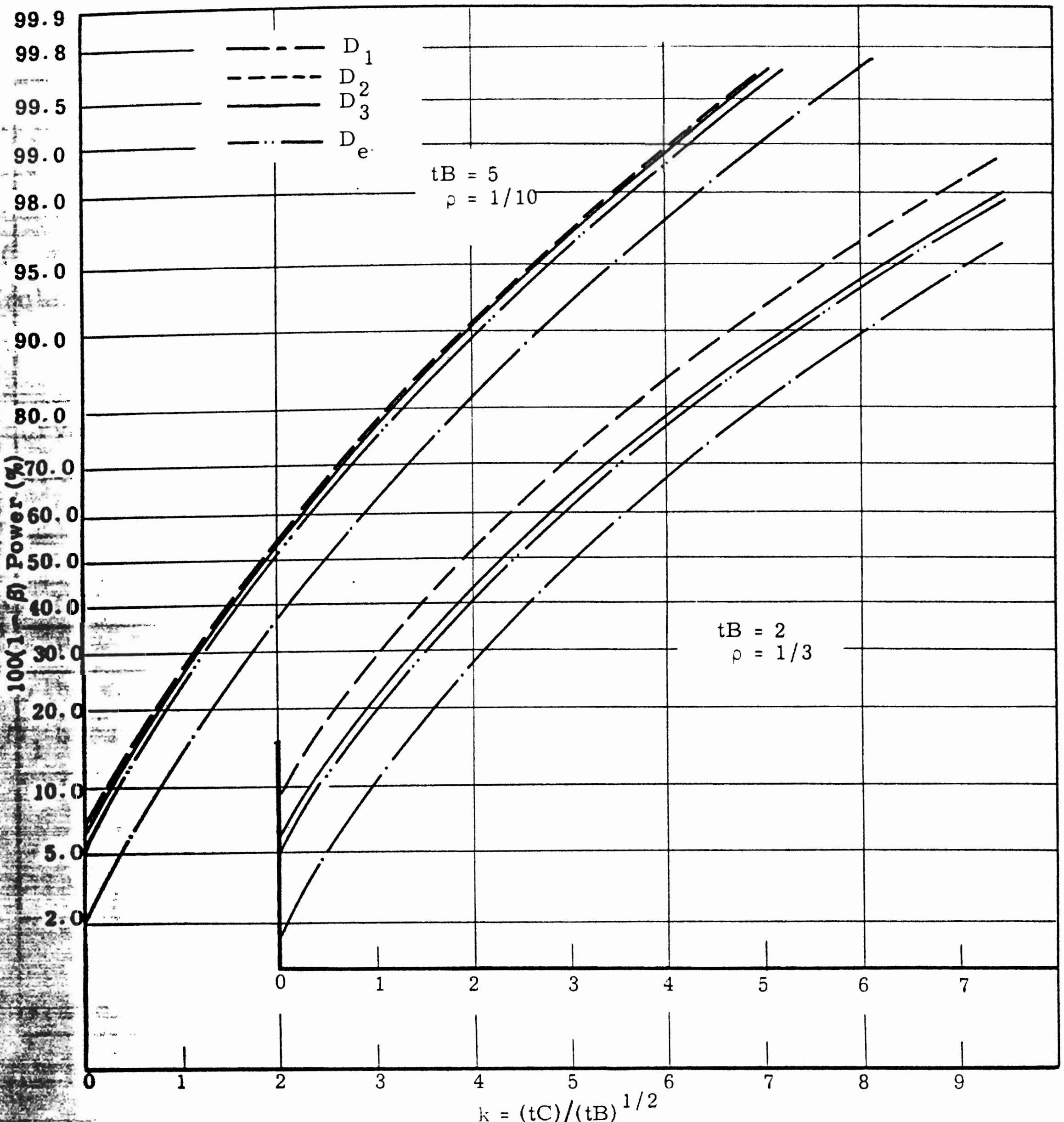

FIGURE 3

Power Surface Sections for Stated 5% Level of Significance Decision Rules D_1, D_2, D_3, and D_e for Selected tB and ρ

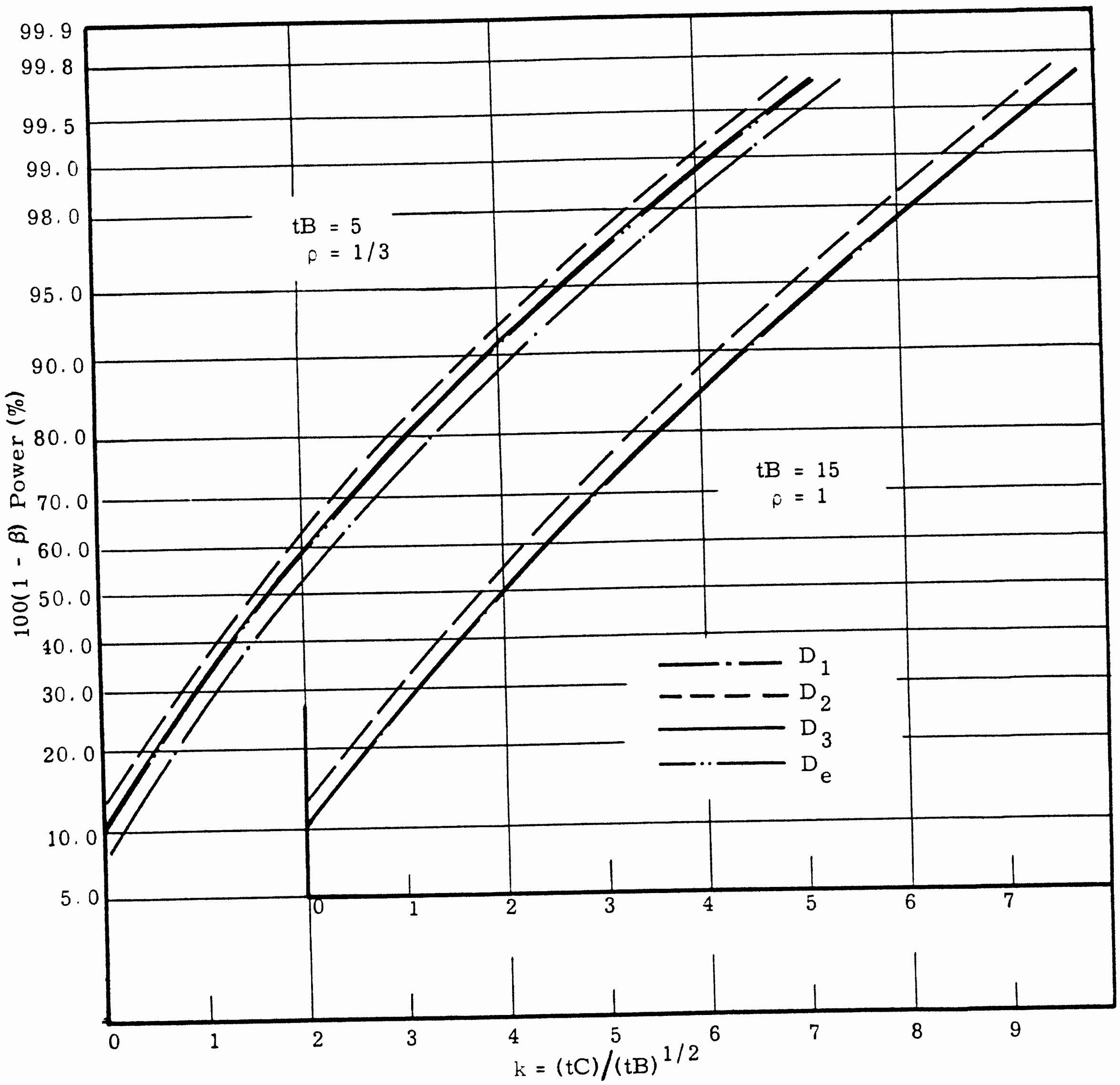

FIGURE 4

Power Surface Sections for Stated 10% Level of Significance Decision Rules D_1, D_2, D_3, and D_e for Selected tB and ρ

Thus, k is the ratio of the expected total net sample count to the square root of the expected total background count in the gross sample count.

The comparisons in Figures 2, 3, and 4 illustrate clearly the relationships among the four power surface sections which are evident for all 180 (ρ, α, tB) triples investigated numerically. First, all four curves are nearly parallel (excluding the region near $C = 0$) for a fixed (ρ, α, tB) triple. This means that the ordering of the powers of the four rules at specific (B_o, C_o) point is the same as the ordering of the exact type I errors at $(B_o, 0)$. Because of the vertical normal probability scale the power increments between the curves do not remain fixed. All curves approach one as C increases. This ordering of the powers points out the importance of keeping the exact type I error close to the stated one. When one of the rules D_1, D_2, and D_3 has a large negative bias in α, the power is depressed over the full range of positive C values. On the other hand, the power is uniformly high for a rule with a large positive bias in α. Since the rule D_3 has the exact type I errors closest to the stated ones over a large range of (ρ, α, tB) triples it is superior to D_1 because of less type I error bias and superior D_2 because of higher power.

Second, with $\rho \leqq 1$ and tB not too small there does not seem to be any advantage in using the uniformly most powerful similar rule D_e unless it is essential that the type I error be known exactly. For $\rho > 1$ and/or tB small the type I errors of the normal approximation rules are quite erratic. For these cases the rule D_e should be used.

Asymptotic Power Surfaces

When $t(B + C)$ and sB increase without bound the power functions of the four decision rules D_1, D_2, D_3, and D_e all approach the same limiting form. This asymptotic power function is derived below for D_3. With trivial changes the same argument applies for D_1 and D_2. It has already been noted above that D_3 and D_e are asymptotically equivalent rules.

Suppose t(B + C) and sB → +∞ in such a manner that k of (27) is a fixed positive constant. The limiting power function of the rule D_3 is

$$\lim_{\substack{sB \to +\infty \\ t(B+C) \to +\infty}} \Pr\{x - \rho y > n_\alpha[\rho(x+y)]^{1/2} \,|\, H_C\} = F[k(1+\rho)^{-1/2} - n_\alpha], \tag{28}$$

where F is the cumulative unit normal distribution function defined by (10).

The proof of (28) is straightforward. Let u and v be chance quantities defined as

$$u = (x - \rho y)(tC)^{-1} \text{ and } v = \rho(x + y)(tC)^{-2}. \tag{29}$$

Since x and y are mutually independent Poisson variables with means t(B + C) and tB/ρ, respectively, from (29) we know

$$\begin{aligned} \text{Ave}(u) &= 1; \\ \text{Ave}(v) &= \rho(tC)^{-1} + (1+\rho)k^{-2} \to (1+\rho)k^{-2}; \\ \text{Var}(u) &= (tC)^{-1} + (1+\rho)k^{-2} \to (1+\rho)k^{-2}; \\ \text{Var}(v) &= \rho(tC)^{-3} + (1+\rho)(ktC)^{-2} \to 0. \end{aligned} \tag{30}$$

The limits in (30) are approached as t(B + C), sB → +∞ consistent with (27). Now u as a linear function of mutually independent Poisson variables which individually in the limit are normally distributed also becomes a normal variable in the limit. From (30) $v \to (1+\rho)k^{-2}$ in probability. By a version of Slutsky's theorem (see [6], p. 254),

$$\begin{aligned} &\lim_{\substack{sB \to +\infty \\ t(B+C) \to +\infty}} \Pr\{x - \rho y > n_\alpha[\rho(x+y)]^{1/2} \,|\, H_C\} \\ &= \lim_{\substack{sB \to +\infty \\ t(B+C) \to +\infty}} \Pr\{u > n_\alpha v^{1/2} \,|\, H_C\} \\ &= \Pr\{u_\infty > n_\alpha(1+\rho)^{1/2}k - 1\}, \end{aligned} \tag{31}$$

where u_∞ is a normal chance variable with mean 1 and variance $(1 + \rho)k^{-2}$. Evaluation of the last probability in (31) in terms of the unit normal cumulative function gives the desired result in (28).

The sets of asymptotic power surface sections for $\alpha = 0.01$, 0.05, and 0.10, respectively, are included in Appendix D. Each set contains curves for $\rho = 0$, 1/10, 1/5, 1/3, 1/2, 1, 2, 3, 5, 10, 25, and 100.

V. CONFIDENCE INTERVAL ESTIMATES

Let x and y be independent Poisson change variables with mean values t(B + C) and sB, respectively. As for the hypothesis testing problem with t and s known there is a one - one correspondence between the possible p.d.f.'s for x and y and the parameter space Ω consisting of all points (B, C) in the closure of the first quadrant of the two-dimensional Cartesian plane. The problem is to construct a confidence interval estimate of C based only on the knowledge of x, y, s, and t.

Two solutions to the confidence interval problem are described in the sequel. The primary distinction between them is the treatment of the unknown background parameter B. The first interval I_1 is the confidence interval formulation of the symmetric form of the hypothesis testing rule D_1. I_1 replaces B by the unbiased estimate y/s. The second interval I_4 is an example of a familiar method of eliminating a nuisance parameter (see[12, 13]) based on joint confidence region estimation of the pair (B, C). The interval I_1 is classical, the method suggested in most texts on radiochemistry. The interval I_4 has not appeared in print before. Both are based on the normal approximation to the Poisson distribution. For small t(B + C) and sB I_4 is superior to I_1 in some aspects to be explained below. To date, no exact, much less best in any sense, confidence interval for C has appeared in the literature so there is no absolute standard against which these normal approximation procedures can be measured.

Intervals Based on Normal Approximation

With the assumption that x and y are normally distributed, an approximate $100(1 - \alpha)\%$ confidence interval for either the net count tC, or the net count rate C, can be constructed using the two unbiased statistics, the sample net count,

$$t\tilde{c} = x - \rho y, \tag{32}$$

and the estimate of the variance σ^2 in (11) of the sample net count,

$$\tilde{\sigma}_1^2 = x + \rho^2 y. \tag{33}$$

The net count form of this confidence interval I_1 is

I_1. Decide that C satisfies

$$x - \rho y - n_{\alpha/2}(x + \rho^2 y)^{1/2} \leqq tC \leqq x - \rho y + n_{\alpha/2}(x + \rho^2 y)^{1/2}. \tag{34}$$

Here, $\rho = t/s$ and $n_{\alpha/2}$ is the $100(1 - \alpha/2)$ percentile of the unit normal distribution (see the discussion leading up to (10)). The rule I_1 is the confidence interval formulation of the symmetric decision rule version of the one-sided rule D_1.

An alternative interval can be constructed using a joint confidence region approach. With x and y independent normally distributed chance variables the chance function

$$\chi^2 = (x - u)^2/u + (y - v)^2/v, \tag{35}$$

where $u = t(B + C)$ and $v = sB$, has a chi-square distribution with two degrees of freedom. Since $C \geqq 0$, the pair (u, v) must satisfy $u \geqq \rho v$. Let χ^2_β be the 100β percentile on this chi-square distribution. Then,

$$\Pr\{\chi^2 \leqq \chi^2_{1-\alpha}\} = 1 - \alpha. \tag{36}$$

independently of u and v. The random set S(x, y) of points in the first quadrant of the uv-plane, which is the intersection of the bracketed statement of (36), and the half-plane $\{u \geqq \rho v\}$ is a $100(1 - \alpha)\%$ confidence region for the true parameter pair (u, v). It can be represented as

$$S(x, y) = \{(u, v) \,\big|\, (x - u)^2/u + (y - v)^2/v \leqq \chi^2_{1-\alpha},\ u \geqq \rho v \geqq 0\}. \tag{37}$$

The set S(x, y) is closed and strictly bounded. As such, any real, continuous function of (u, v) takes on a maximum and minimum value on S. (See Appendix C for the mathematics to justify this argument.) The expected net sample count, $tC = u - \rho v$, is such a function. With t and s fixed the interval (L, U), where

$$L = \min_{(u,v)\epsilon S(x,y)} (u - \rho v),$$

$$\text{and} \quad U = \max_{(u,v)\epsilon S(x,y)} (u - \rho v), \tag{38}$$

is a confidence interval for the true expected net count with confidence coefficient greater than $(1 - \alpha)$. The equivalent interval for count rate is (L/t, U/t). The calculation of L and U given x, y, ρ, and α is quite tedious. L and U must be approximated by numerical means as they are functions of the roots of an irreducible sixth degree polynomial.

Instead of tabling the conservative interval endpoints in (38), the normal percentile point $n_{\alpha/2}$ was substituted for $(\chi^2_{1-\alpha})^{1/2}$ to give the confidence interval

I_4. Decide that C satisfies

$$L_4 \leqq tC \leqq U_4. \tag{39}$$

That is, L_4 and U_4 are defined as L and U in (38) with the set $S(x, y)$ of (37) modified by replacing $\chi_{1-\alpha}$ with $n_{\alpha/2}$. Tables of the rule I_4 are given in Appendix D. Because the set $S(x, y)$ is contained in the half-plane $\{u \geqq \rho v\}$ both L_4 and U_4 are necessarily nonnegative. For comparison of expected lengths the rule I_1 was also forced to give a nonnegative interval by replacing each endpoint with the maximum of the endpoint and zero.

To compare the performance of the two confidence interval rules I_1 and I_4 the exact probability of covering the true expected net sample count and the expected length of the interval were computed for each rule at all 972 possible combinations of the following parameters:

$$tC = 0, 1, 2, 3, 5, 10, 20, 30, 50;$$
$$tB = 1, 2, 3, 5, 10, 20;$$
$$\rho = 1/10,\ 1/5,\ 1/3,\ 1/2,\ 1, 2;$$
$$1 - \alpha = 0.90,\ 0.95,\ 0.99$$

These calculations generally show that I_1 gives confidence intervals of shorter expected length, and that I_4 has an exact confidence coefficient closer to the stated one. Also, that for I_4 the exact coefficient errs on the conservative side (is greater than the stated one), while for I_1 the exact coefficient errs on both sides of the stated one.

Table II gives a typical comparison of the two rules for a stated confidence coefficient of 0.90 and $\rho = 1/5$. The four columns of the table give exact confidence coefficients and expected lengths for an expected background count of 2 and various expected net sample counts. Rule I_1 exact coefficients vary between 0.985 and 0.867 while those for I_4 between 0.938 and 0.906. For larger background counts the discrepancies between the rules are not as great. Examination of similar tables for the range of $(1 - \alpha,\ \rho)$ pairs investigated numerically indicate that the major differences between the two rules occur for expected background count of less than ten.

TABLE II

COMPARISON OF CONFIDENCE INTERVAL RULES I_1 AND I_4 FOR STATED 0.90 CONFIDENCE COEFFICIENT, $\rho = 1/5$, AND EXPECTED BACKGROUND tB = 2

	Probability of Covering True Expected Net Sample Count tC		Expected Length of Confidence Interval	
tC	I_1	I_4	I_1	I_4
0	.985	.938	2.7	4.0
1	.869	.906	4.0	5.3
2	.871	.909	5.2	6.5
3	.867	.908	6.4	7.5
5	.878	.913	8.3	9.1
10	.889	.911	11.4	11.9
20	.894	.907	15.5	15.9
30	.894	.906	18.7	19.1
50	.896	.906	23.8	24.2

Appendix D lists the confidence interval endpoints L_4 and U_4 for rule I_4 for all combinations of $1 - \alpha, \rho, x$, and y compatible with an expected background count of not more than ten during the recording of the sample count. For x and/or y beyond the range of the table, the simple rule I_1 can be used. Also in Appendix D the complete set of exact confidence coefficient expected length tables are listed for both rules I_1 and I_4.

Asymptotic Expected Lengths

When t(B + C) and sB increase without bound the expected lengths of the two confidence interval rules I_1 and I_4 approach the same limiting form. Suppose t(B + C), sB $\rightarrow +\infty$ in such a manner that the standardized expected net sample total count k is constant. This is the same limiting process discussed for hypothesis testing, (see (28)). Let E be the expected length of either I_1 or I_4. Then,

$$\lim_{\substack{t(B+C) \rightarrow +\infty \\ sB \rightarrow +\infty}} (E/ct) = 2n_{\alpha/2}(1 + \rho)^{1/2}/k. \tag{40}$$

es

Clearly, when the expected net sample total count $Ct \to +\infty$, the length of the confidence interval must also approach $+\infty$. The result (40) shows that the ratio of these two quantities approaches a limit. Figure D.1 in Appendix D is a plot of this limiting ratio as a function of ρ and k. The ratio is normalized by division with $2n_{\alpha/2}$ so a single graph suffices for all confidence coefficients.

The proof of (40) for I_1 is straightforward. Let u be the chance quantity defined as

$$u = (x + \rho^2 y)^{1/2}/(tC). \tag{41}$$

Rewrite (41) to give

$$\left|u^2 - \frac{1+\rho}{k^2}\right| = \left|\frac{1}{k^2}\,\frac{x - t(B+C)}{t(B+C)} + \frac{1}{tC}\,\frac{x}{t(B+C)} + \frac{\rho}{k^2}\,\frac{y - sB}{sB}\right|. \tag{42}$$

Using the triangle inequality and taking expectations of both sides of (42)

$$\begin{aligned} \text{Ave}\left|u^2 - \frac{1+\rho}{k^2}\right| &\leqq \frac{1}{k^2}\,\text{Ave}\frac{|x - t(B+C)|}{t(B+C)} + \frac{1}{tC}\,\text{Ave}\frac{|x|}{t(B+C)} + \frac{\rho}{k^2}\,\text{Ave}\frac{|y - sB|}{sB} \\ &\leqq \frac{1}{k^2}\left[\frac{1}{t(B+C}\right]^{1/2} + \frac{1}{tC} + \frac{\rho}{k^2}\left[\frac{1}{sB}\right]^{1/2}. \end{aligned} \tag{43}$$

From (43) $u^2 - \dfrac{1+\rho}{k^2} \to 0$ in mean of order one when $t(B + C)$ and $sB \to +\infty$.

The inequality,

$$\left|u - \frac{(1+\rho)^{1/2}}{k}\right| \leqq \frac{k}{(1+\rho)^{1/2}}\left|u^2 - \frac{1+\rho}{k^2}\right|, \tag{44}$$

(note $u \geqq 0$) gives $u \to (1 + \rho)^{1/2}/k$ in mean of order one. Now (40) follows since by definition

$$\text{Ave}\,(u) = E/2n_{\alpha/2}\,Ct \tag{45}$$

for rule I_1.

VI. THE IMPORTANCE OF GOOD BACKGROUND ESTIMATION

1 In both the hypothesis testing and confidence interval problems discussed in Sections IV and V the goodness of the inference about the unknown net sample count rate C improves with more precise information about the background count rate B. For a fixed expected total background count tB contribution to the sample gross count, the importance of the background estimate can be investigated as a function of the total expected net sample count tC and the sample to background counting time ρ by reorganizing the results presented in Sections IV and V.

Hypothesis Testing

Consider the family of power surface sections for a given decision rule D with tB and α fixed and ρ varying over a set of non-negative values. The curve $\rho = 0$ represents the limiting case when the expected background count is known exactly. Positive ρ curves represent different degrees of information about the background where information increases as ρ decreases. Three of these families of power curves for rule D_3 and (tB, α) pairs of (10, 0.10), (2, 0.05), and (50, 0.01) are plotted in Figures 5, 6, and 7 respectively. The limiting families for α = 0.01, 0.05, and 0.10 applicable for large t(B + C) and sB, and all four decision rules D_1, D_2, D_3, and D_e are plotted in
+∞ Appendix B. The effect on power of increasing the length of the counting period for the background estimate is clearly seen in these plots. The power is very low compared to the optimum curve ($\rho = 0$) if $\rho > 1$. Good power demands at least $\rho = 1$ (i.e., counting the background at least as long as the sample). The economics of the particular application would undoubtedly determine what ρ in the interval $0 < \rho \leq 1$ is optimum. Certainly the potential increase in power hardly warrants a $\rho < 1/5$. A reasonable rule of thumb seems to call for a ρ in the range 1 to 1/3.

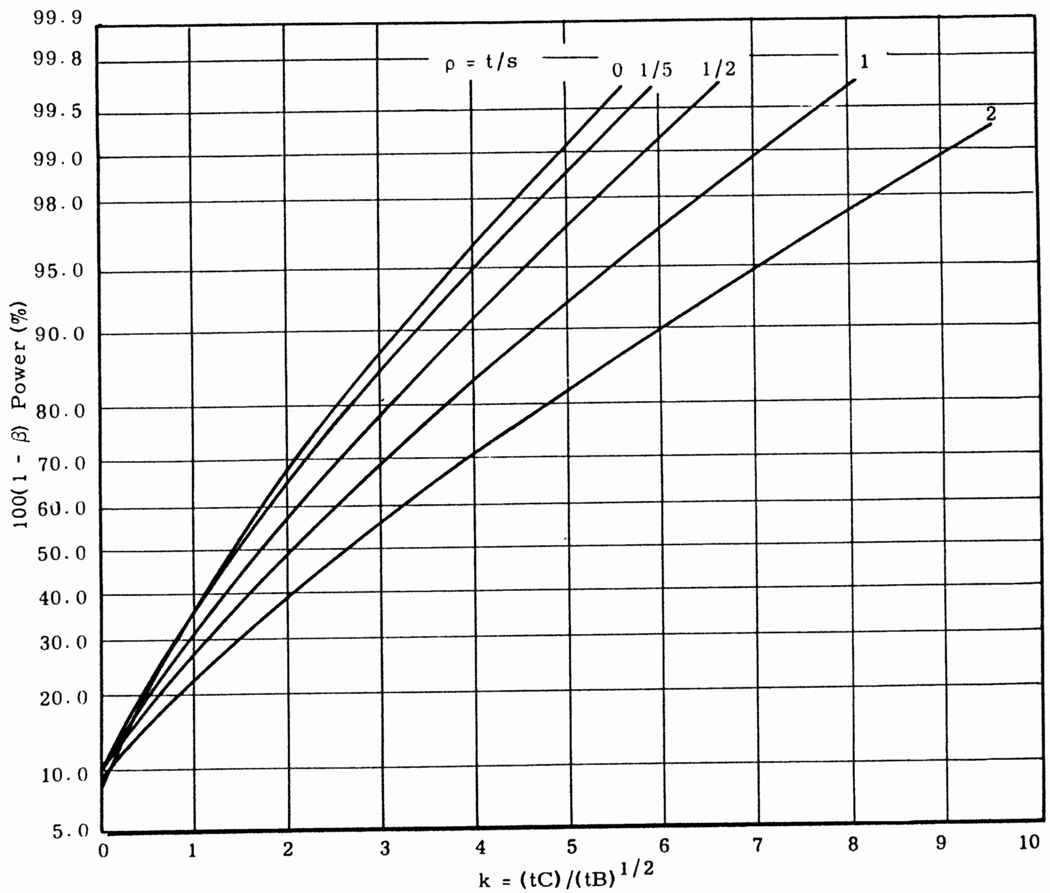

FIGURE 5

Stated 10% Level of Significance Power Surface Sections for Decision Rule D_3 and Expected Background tB = 10 Counts

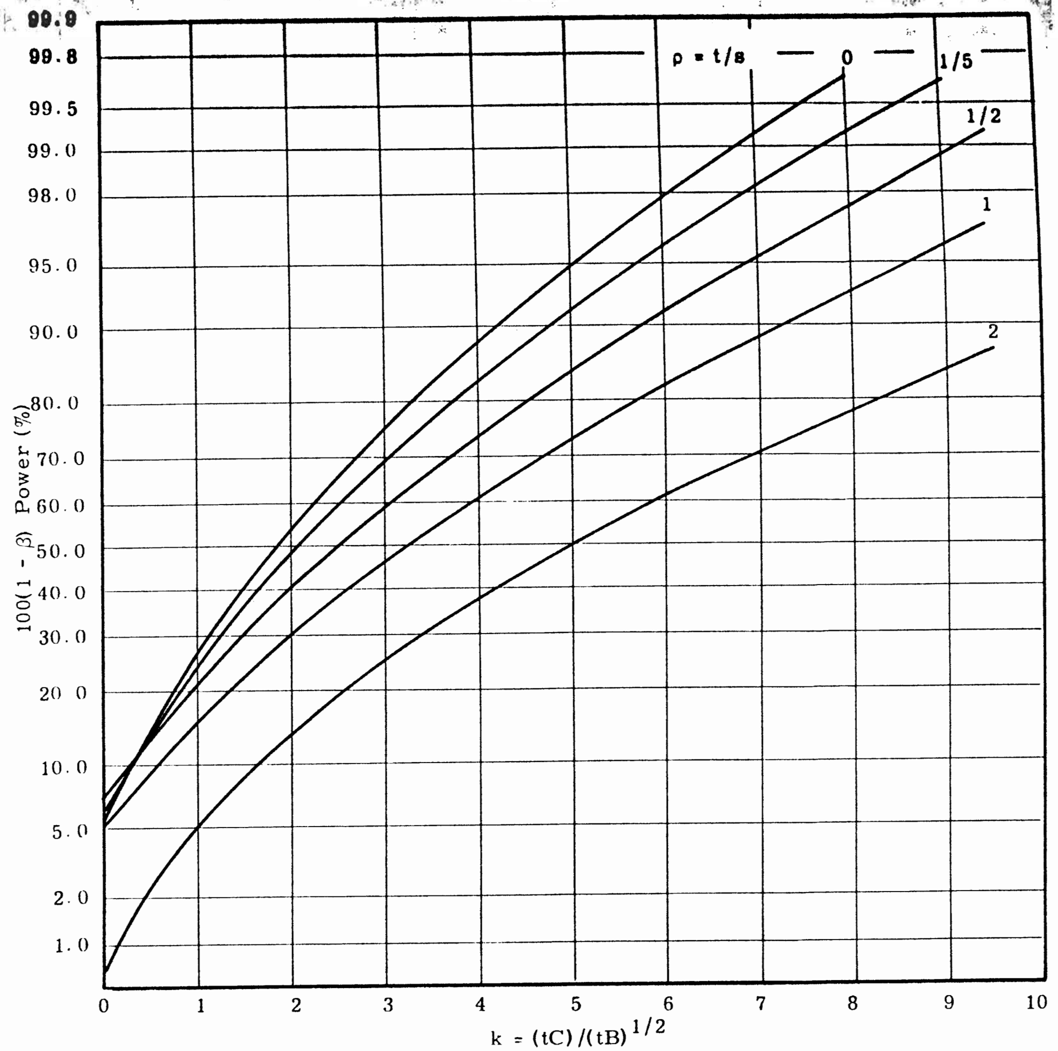

FIGURE 6

Stated 5% Level of Significance Power Surface Sections for Decision Rule D_3 and Expected Background tB = 2 Counts

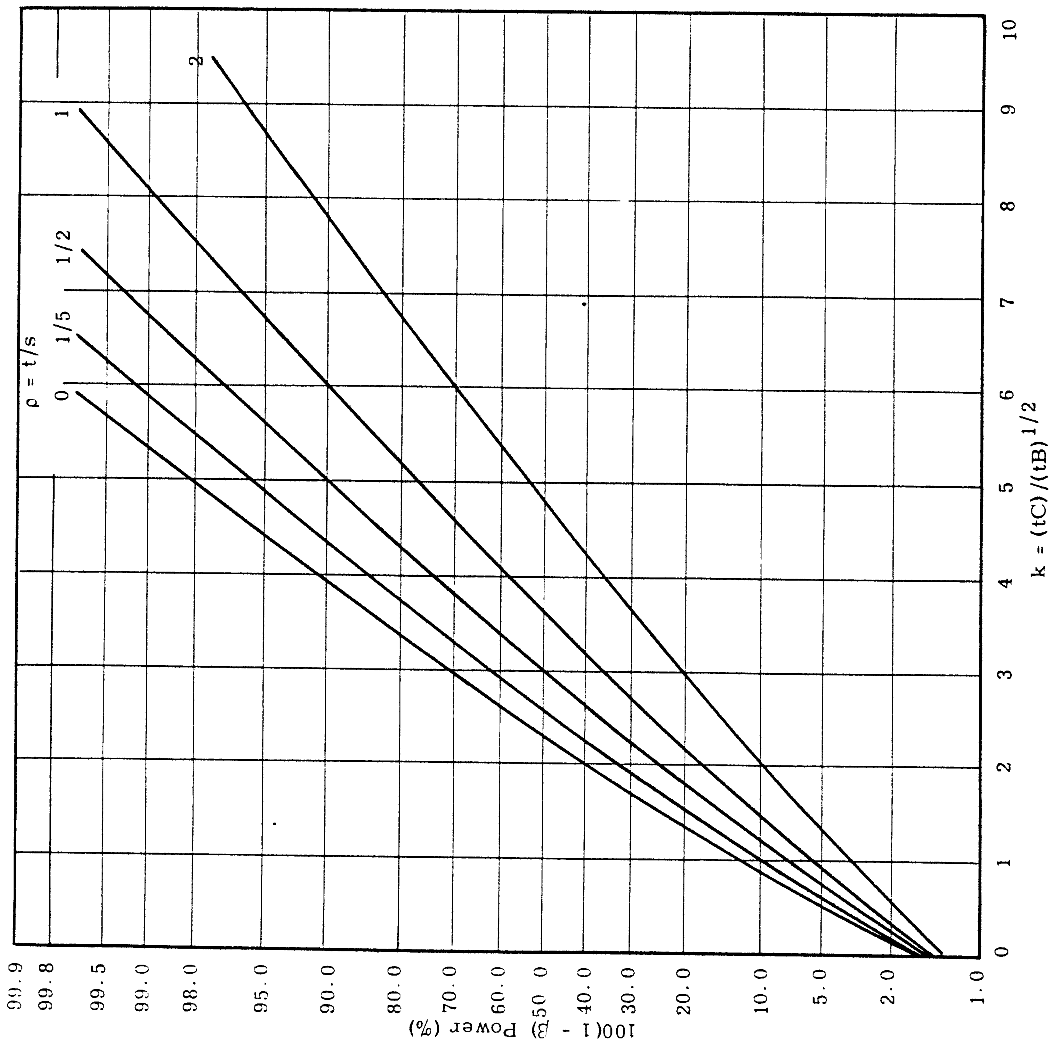

FIGURE 7

Stated 1% Level of Significance Power Surface Sections for Decision Rule D_3 and Expected Background tB = 50 Counts

A more intelligent decision can be made if good power is required for a particular alternative or short range of alternatives. For example, in the limiting power curve families the maximum difference between the power curve for $\rho = 0$ and curves for $0 < \rho \leqq 1$ is for k near 2. The exact position of k shifts slightly with α and ρ. For good power near $k = 2$, $\rho < 1$ is needed, but at least on an absolute scale for large k say $k > 5$, $\rho = 1$ will suffice.

The above discussion applies when the sample count has already been taken or, at least, the decision has been made as to how long the sample is to be counted. A background count must now be taken which is of commensurate precision so that most of the information in the sample will be extracted. A different problem arises if a fixed time period is available in which both sample and background counts must be taken. In the asymptotic case when both B and C increase while $C/B^{1/2}$ remains bounded, splitting the allotted time equally to sample and background is optimum; i.e., $\rho = 1$ is optimum. This follows directly from the fact that the argument of the cumulative unit normal distribution of the limiting power function of (28) can be written as

$$\frac{k}{(1+\rho)}\,1/2 - n_\alpha = \left(\frac{st}{s+t}\right)^{1/2} \frac{C}{B^{1/2}} - n_\alpha. \qquad (46)$$

With $s + t = M$, a constant, (46) is a function of s and t only through the product st, which is a maximum for $s = t = M/2$. No formal proof is offered of the optimum allocation for the finite expected count exact theory. Numerical checks using the power surface sections for rule D_3 in Appendix B suggest that if the optimum is not $\rho = 1$, it is close to it. Figure 8 shows six plots of power versus ρ with $t + s$ fixed, three being limiting theory and three exact. Each curve is identified by a k value or k and tB values when $\rho = 1$. These curves are typical in that the optimum is quite loosely defined. From a practical standpoint, any ρ in the range $1/2 \leqq \rho \leqq 3/2$ gives power close to the maximum.

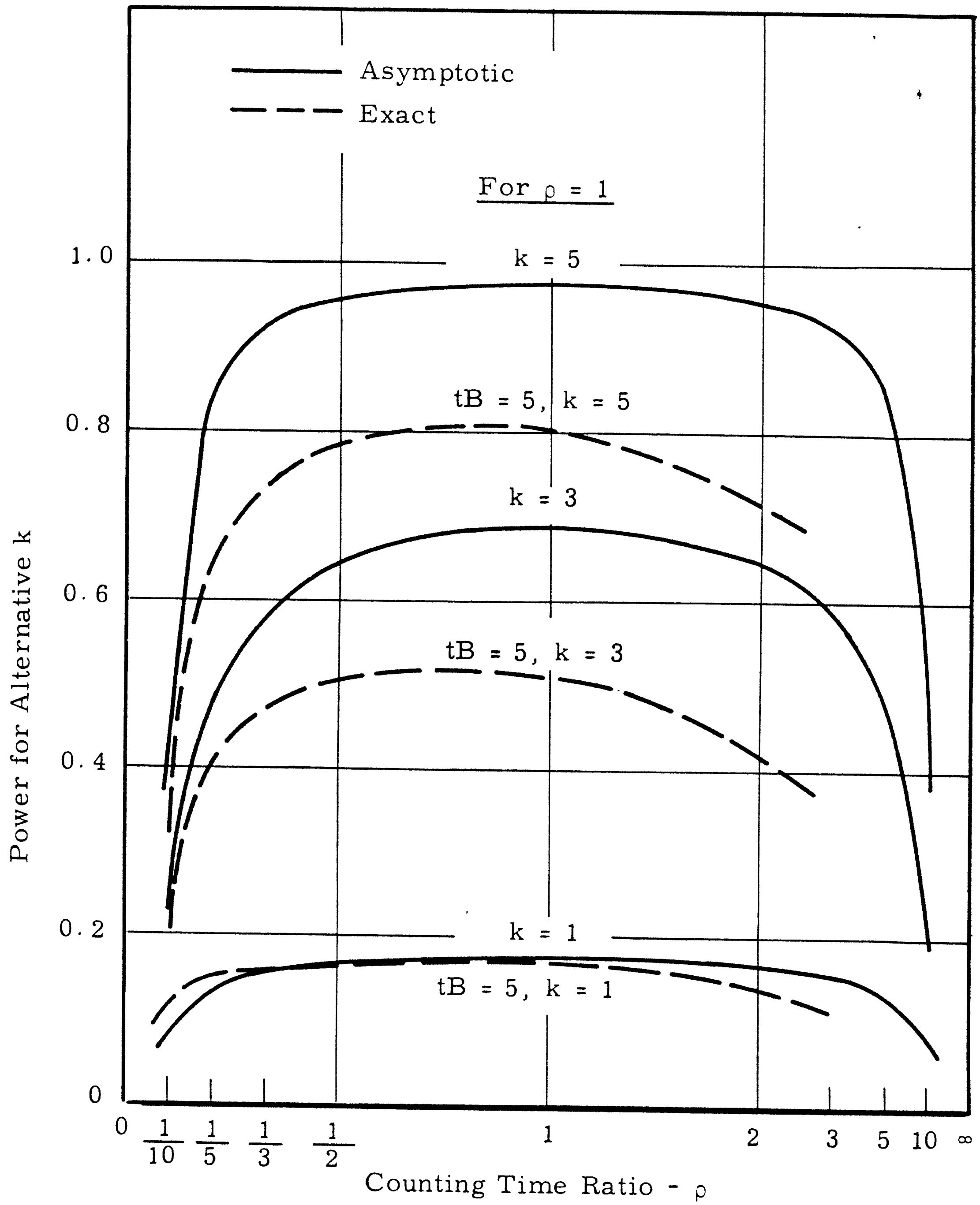

FIGURE 8

Power of Stated 5% Level of Significance Decision Rule D_3 as a Function of $\rho = t/s$ for Fixed Total Counting Time $t + s = 2k^2B/C$

Confidence Intervals

In a manner similar to hypothesis testing, families of expected confidence interval length can be plotted for fixed confidence coefficient $(1 - \alpha)$ and background expected count tB, and a range of ρ values including $\rho = 0$. Three such families of curves for (tB, $1 - \alpha$) pairs of (2, 0.95), (10, 0.90) and (20, 0.99) are plotted in Figures 9, 10, and 11 respectively. Since the rule I_4 was suggested for an expected background count of less than 10 and I_1 for greater than 10, Figure 9 is for I_4. Figure 10 as the transition point includes both I_1 and I_4 and Figure 11 is for I_1. The limiting family for large t(B + C) and sB is plotted in Figure D.1 of Appendix D. These plots are normalized to include all confidence coefficients.

The conclusions for all these plots are similar to those for hypothesis testing. Here, improvement in the rule is measured by the incremental shortening of the expected length function in place of the incremental increasing of the power function. Clearly, short expected length demands $\rho \leqq 1$. The exact value of $0 < \rho \leqq 1$ would depend upon the application. A good rule of thumb seems to be, as above, a ρ in the range 1 to 1/3 depending on the economics of the particular application.

The picture is particularly clear for the limiting expected length function of (40). For a given sample and background the ratio of the expected length for a given ρ to the shortest expected length with $\rho = 0$ is just $(1 + \rho)^{1/2}$.

With a fixed total counting time for both sample and background, an argument identical to that for hypothesis testing shows that $\rho = 1$ gives the shortest expected length in the asymptotic case. Again numerical checks suggest a similar result for the exact theory.

At first glance the result of $\rho = 1$ being optimum may seem to contradict the familiar rule based on counting rates (see[16], p. 269, prob. 6). Specifically, the rule is to select ρ so that

$$\rho^2 = (C + B) \Big/ B \qquad (47)$$

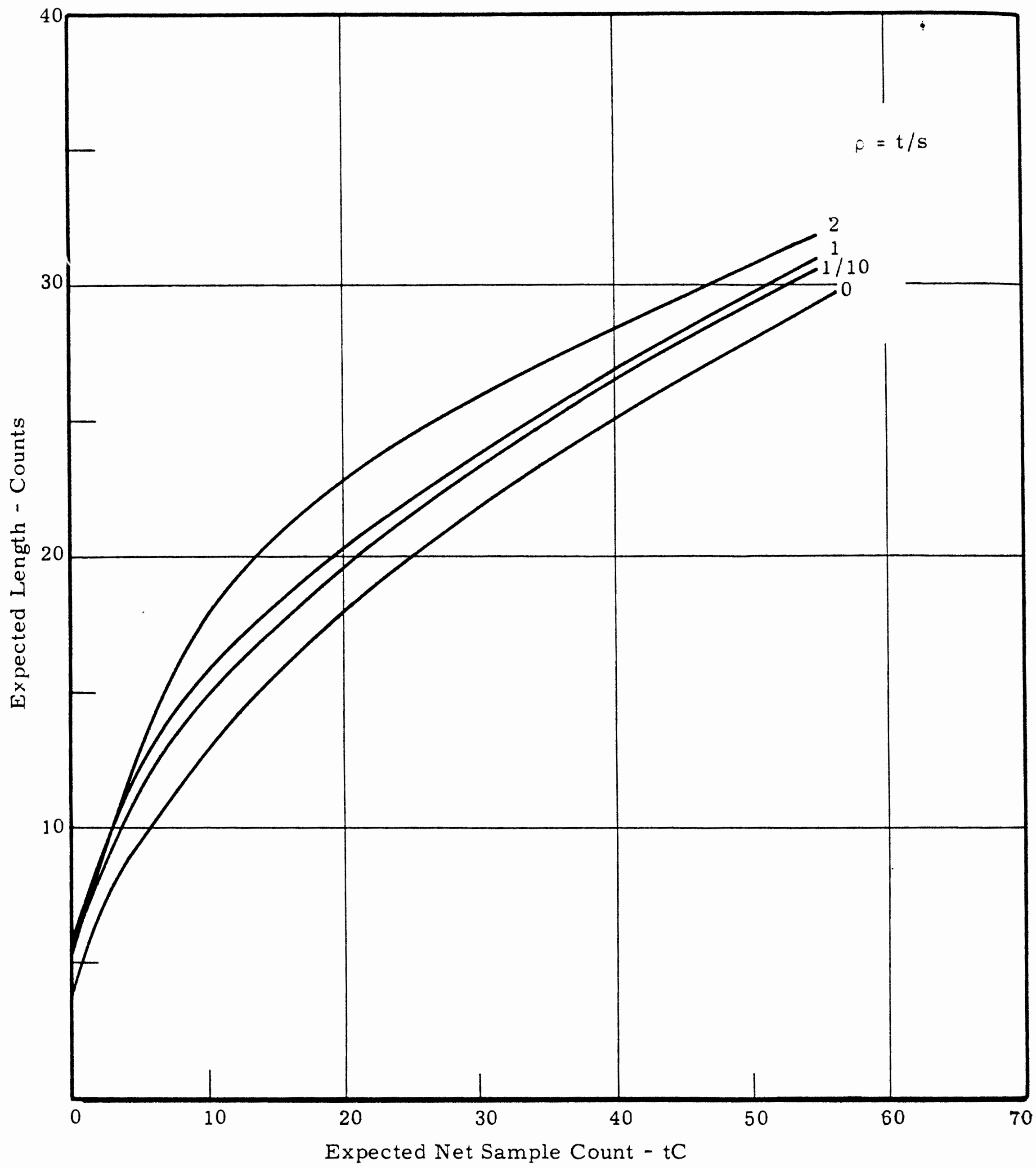

FIGURE 9

Stated 95% Confidence Interval Expected Length
for Confidence Interval Rule I4 for Expected Background
tB = 2 Counts

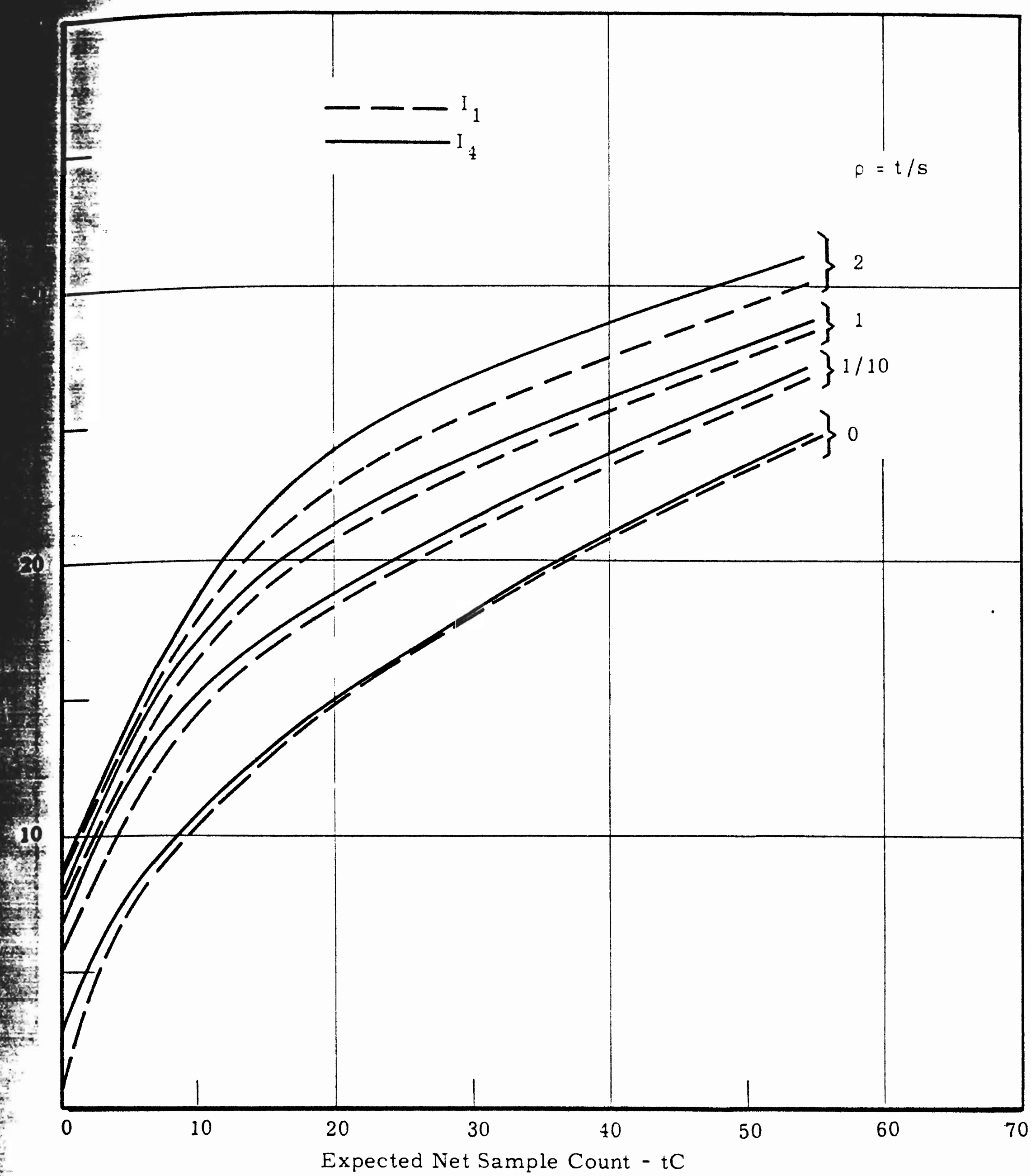

FIGURE 10

Stated 90% Confidence Interval Expected Length for Confidence Interval Rules I_1 and I_4 for Expected Background tB = 10 Counts

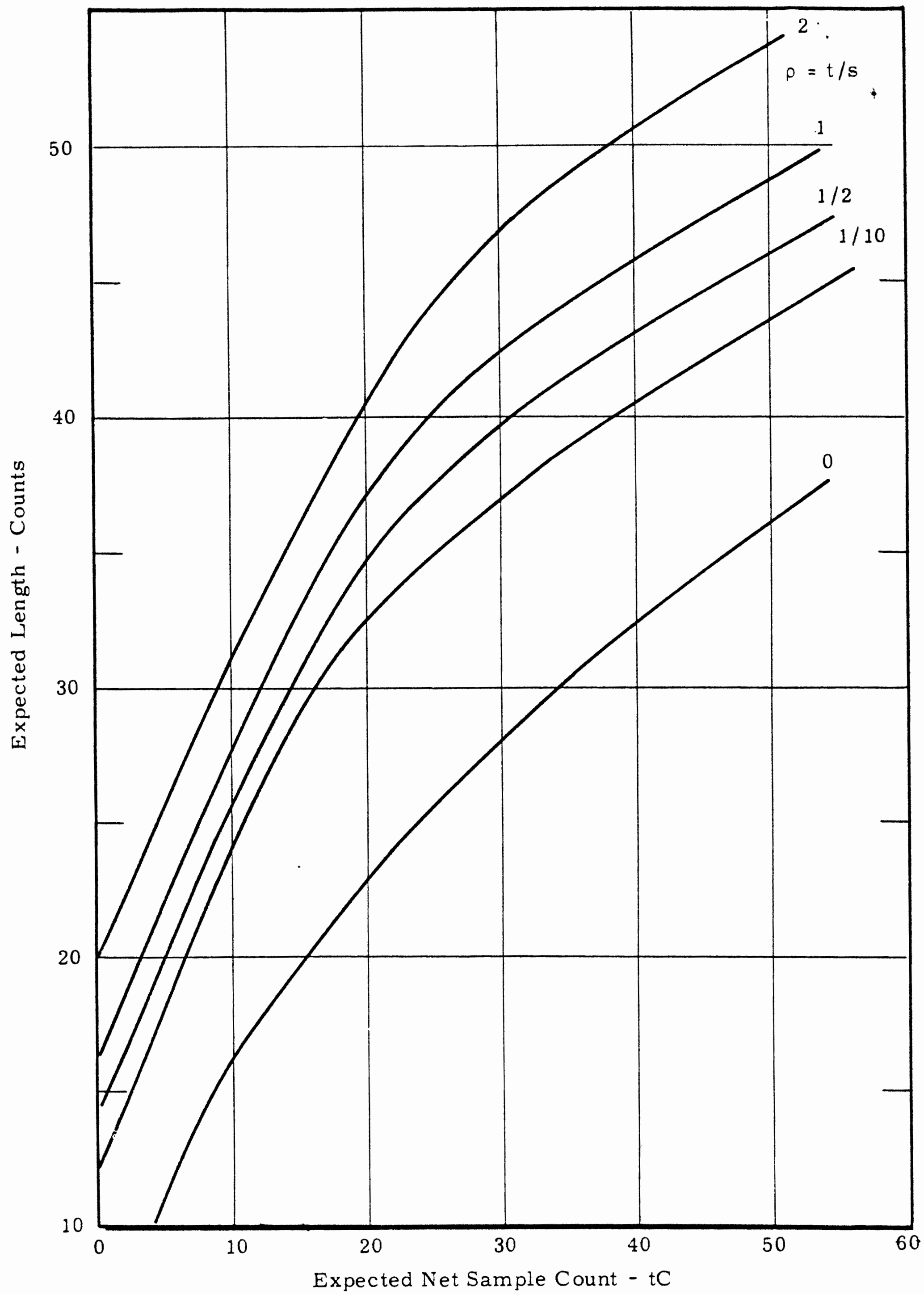

FIGURE 11

Stated 99% Confidence Interval Expected Length
for Confidence Interval Rule I_1 for Expected Background
tB = 20 Counts

fixed total counting time must be allocated to both sample and round. The rationale behind (47) is that such a ρ minimizes the nce of $\tilde{c}$. When the means t(C + B) and sB are small both x and Poisson variables have distinctly non-normal distributions. As minimizing the variance of $\tilde{c}$ does not necessarily maximize the of a decision rule and/or give the shortest confidence interval th. When both means are large, x and y are approximately nor- and minimizing the variance of $\tilde{c}$ is the correct procedure. How- the limit of the count rate ratio in (47) is one when t(B + C), $\to \infty$ so that $k = (tC)/(tB)^{1/2}$ remains fixed. For the cases considered his paper the above rule in the limit agrees with our recommenda- ns.

ltiple Use of Background Estimates

Frequently in practice, the same background estimate may be d many times to correct sample total counts to net counts. For mple, a radiochemical laboratory may check the background of a rticular counting instrument once or twice daily. Either the latest ckground estimate or a long term average estimate is used with all mples analyzed on the instrument until another background check is de.

The average value characteristics of decision rules, both ypothesis testing and confidence interval, when multiple use is made background estimates, are the same as those discussed above where n independent background estimate is provided for every sample. If sample analysis is selected at random, the parameter triple (k, tB, ρ) defines the power of a hypothesis testing rule and/or the expected length of a confidence interval rule in the manner specified in Sections IV and V. Multiple use of background estimates does not influence the average behavior of rules on randomly selected samples or the long term behavior averaged over background estimates; but it does increase the variability of the characteristics of rules about the average value, since the decisions tend to be blocked on background.

To illustrate the situation consider a fixed number T of samples that are counted and background estimates supplied for each. The samples are grouped into R blocks of size $M_1, M_2, \ldots, M_r$. All sample counts for a given block are paired with the same background estimate. Algebraic complexity hardly warrants a general treatment so it is assumed that all T sample counts and R background counts are mutually independent Poisson chance variables, the sample counts having expected value $t(B + C)$ and the background counts expected value sB. Let x_{vj} and y_v be the j^{th} sample count and the background count, respectively, for v^{th} block. The D_3 decision rule for the vj^{th} sample is

Decide H_C is true if $x_{vj} - y_v > n_\alpha[\rho(x_{vj} + y_v)]^{1/2}$;

Otherwise, decide H_o is true.

The I_4 rule is

Decide that C satisfies $L_4 \leqq tC \leqq U_4$,

where L_4 and U_4 are defined in terms of the set $S(x_{vj}, y_v)$ in Section V. Let P(y) be the conditional probability that the decision "H_C is true" is made given that $y_v = y$. Let E(y) be the conditional expected length of the confidence interval given that $y_v = y$. Symbolically,

$$\begin{aligned} P(y) &= \Pr\{x_{vj} - \rho y_v > n_\alpha[\rho(x_{vj} + y_v)]^{1/2} \mid y_v = y\}, \\ E(y) &= \text{Ave}\{U_4 - L_4 \mid y_v = y\}. \end{aligned} \tag{48}$$

Because the T sample counts are identically distributed, P(y) and E(y) do not depend upon v and/or j. Let n_v be the number of times decision "H_C is true" is made for the v^{th} block. Let e_v be the sum of the lengths of confidence intervals for the v^{th} block. Let

$$n = \sum_{v=1}^{R} n_v \text{ and } e = \sum_{v=1}^{R} e_v. \tag{49}$$

Then, n/T is the fraction of correct decisions by rule D_3 for the entire set of T samples. And, e/T is the average length of confidence intervals I_4 for the entire set of T samples. Each n_v and e_v conditional on fixed background count y is the sum of M_v mutually independent chance variables with average values P(y) and E(y), respectively. Using conditional probability calculus

$$\begin{aligned} \text{Ave}\,(n/T) &= (1/T) \sum_{v=1}^{R} \text{Ave}\,[\text{Ave}(n_v | y)] \\ &= (1/T) \sum_{v=1}^{R} \text{Ave}\,[M_v\, P(y)] \\ &= \text{Ave}\,[P(y)] = 1 - \beta_3(C)\,. \end{aligned} \tag{50}$$

Here, $1 - \beta_3(C)$ is the power of rule D_3 for alternative hypothesis H_C discussed in Section IV. Similarly

$$\text{Ave}\,(e/T) = E_4, \tag{51}$$

where E_4 is the expected length of confidence interval rule I_4 discussed in Section V.

Each n_v conditional on fixed background count y is a binomial chance variable on M_v trials with success probability P(y). Using this fact and the conditional variance formula (see Reference[14], p. 65) the variance of n/T is

$$\mathrm{Var}(n/T) = (1/T)^2 \sum_{v=1}^{R} \mathrm{Var}(n_v)$$

$$= (1/T)^2 \sum_{v=1}^{R} \{\mathrm{Ave}\,[\mathrm{Var}(n_v|y)] + \mathrm{Var}\,[\mathrm{Ave}(n_v|y)]\}$$

$$= (1/T)^2 \sum_{v=1}^{R} \{M_v\, \mathrm{Ave}(P(y)[1 - P(y)]) + M_v^2\, \mathrm{Var}[P(y)]\} \qquad (52)$$

$$= (1/T)\, \mathrm{Ave}\,\{P(y)[1 - P(y)]\} + (1/T)^2 \sum_{v=1}^{R} M_v^2\, \mathrm{Var}\,[P(y)].$$

$$= (1/T)\, \mathrm{Ave}\,[P(y)]\, \mathrm{Ave}\,[1 - P(y)] + (1/T)^2 \sum_{v=1}^{R} M_v\,(M_v - 1)\, \mathrm{Var}\,[P(y)]$$

The last line of (52) is obtained by adding $(1/T)\,\mathrm{Var}\,[P(y)]$ to the first term and subtracting the same from the second term of the line above. A similar argument gives

$$\mathrm{Var}(e/T) = (1/T)\, \mathrm{Var}\,(U_4 - L_4) + (1/T)^2 \sum_{v=1}^{R} M_v(M_v - 1)\, \mathrm{Var}\,[E(y)]. \qquad (53)$$

Equation (50) and (51) show that the average value properties of the rules D_3 and I_4 are independent of the number R of background estimates used for and of the way in which these R estimates are allocated to the T samples. It must be remembered here that the averaging is over not only the population of sample counts for an expected total count of t(B + C) but also over the population of background counts for a mean total count of sB. In the analysis of a single set of T samples the fraction of correct decisions n/T and the average confidence interval length e/T will in general deviate from their average values in (50) and (51) respectively. The deviation is measured in a mean square average sense by the variances in (52) and (53). Both variances have a first term which is the variance if an independent background is supplied for each sample. The second term in each case is the additional variance because backgrounds are used for more than one sample. This term is a product of two factors which measure the contribution due to the variance of the background distribution and the way this distribution is sampled (block pattern). The second factor is a minimum of zero for $M_v = 1(R = T)$ and a maximum of $(T - 1)/T$ for $M_1 = T(R = 1)$.

The above example illustrates that the stability of the frequency of correct decisions for hypothesis testing and of the mean confidence interval length will decrease as the size of the blocks with common background correction increases. This effect can be offset by improving the precision of the background estimate.

An interesting and important question (which is not pursued here) is the relative economics associated with various analysis schedules obtained by varying background estimation precision and block size in the light of sample and background analysis cost and the importance of correct decisions.

The exact calculation of the variances in (52) and (53) for Poisson chance variables with finite mean values must be done numerically. Techniques such as those used to calculate $1 - \beta_3(C)$ and E_4 could be used for this purpose. The asymptotic variances can be calculated using tables of bivariate normal distribution probabilities.[15] Asymptotic results are included below for hypothesis testing. An argument similar to that following (24) shows that

$$\lim_{\substack{sB \to +\infty \\ t(B + C) \to +\infty}} \text{Ave}\,[P^2(y)] = \text{BVN}\,[k(1 + \rho)^{-1/2} - n_\alpha,\ \rho/(1 + \rho)] \tag{54}$$

where the limit is taken with $k = (tC)/(tB)^{1/2}$ fixed. In (54) BVN(h, r) is the joint probability that $u \leq h$ and $v \leq h$ where (u, v) has a bivariate normal distribution with means zero, variances one, and correlation coefficient r. With the results in (54) and (24) the limiting form of Var (n/T), say LIMV(n/t), satisfies

$$\mathrm{LIMV}(n/T) = \lim_{\substack{sB \to +\infty \\ t(B + C) \to +\infty}} \mathrm{Var}\,(n/T)$$

$$= (1/T)F(h)\,[1 - F(h)] + (1/T)^2 \sum_{v=1}^{R} M_v(M_v - 1)\,[\mathrm{BVN}(h, r) - F^2(h)] \qquad (55)$$

where $h = k(1 + \rho)^{-1/2} - n_\alpha$, $r = \rho/(1 + \rho)$ and F is the cumulative normal distribution function defined in (10). An interesting property of LIMV(n/T) is that it is a maximum for $h = 0$. Clearly, the first term is a maximum there for $F(0) = 1/2$. For the second term

$$\frac{d}{dh}[\mathrm{BVN}\,(h, r) - F^2(h)] = 2\{F[(\frac{1 - r}{1 + r})^{1/2} h] - F(h)\}\frac{d}{dh}F(h). \qquad (56)$$

This derivative is positive for $h < 0$, zero for $h = 0$ and negative for $h > 0$, which implies the maximum property for $h = 0$.

For fixed ρ maximizing LIMV(n/T) over k gives

$$\underset{k}{\mathrm{Max}}\ \mathrm{LIMV}(n/T) = (1/4T) + (1/T)^2 \sum_{v=1}^{R} M_v(M_v - 1)\,[\mathrm{BVN}(0, r) - 1/4]. \qquad (57)$$

Another way of looking at the location of the maximum is that LIMV(n/T) is a maximum for the alternative k for which the power of the asymptotic test is 1/2. This alternative is $k = n_\alpha(1 + \rho)^{1/2}$ or $tC = n_\alpha(1 + \rho)^{1/2}(tB)^{1/2}$.

For the simplest case where each background is used the same number of times ($M_v \equiv M$) the max function of (57) reduces to

$$\underset{k}{\mathrm{Max}}\ \mathrm{LIMV}(n/T) = (1/4T) + [(M - 1)/T][\mathrm{BVN}(0, r) - 1/4]. \qquad (58)$$

The normalized form of (58) [the ratio of (58) for arbitrary M to 1/4T, (58) with M = 1] is plotted in Figure 12. The family of curves for various M values clearly shows that to keep LIMV(n/T) small as M increases ρ must be pushed down toward zero.

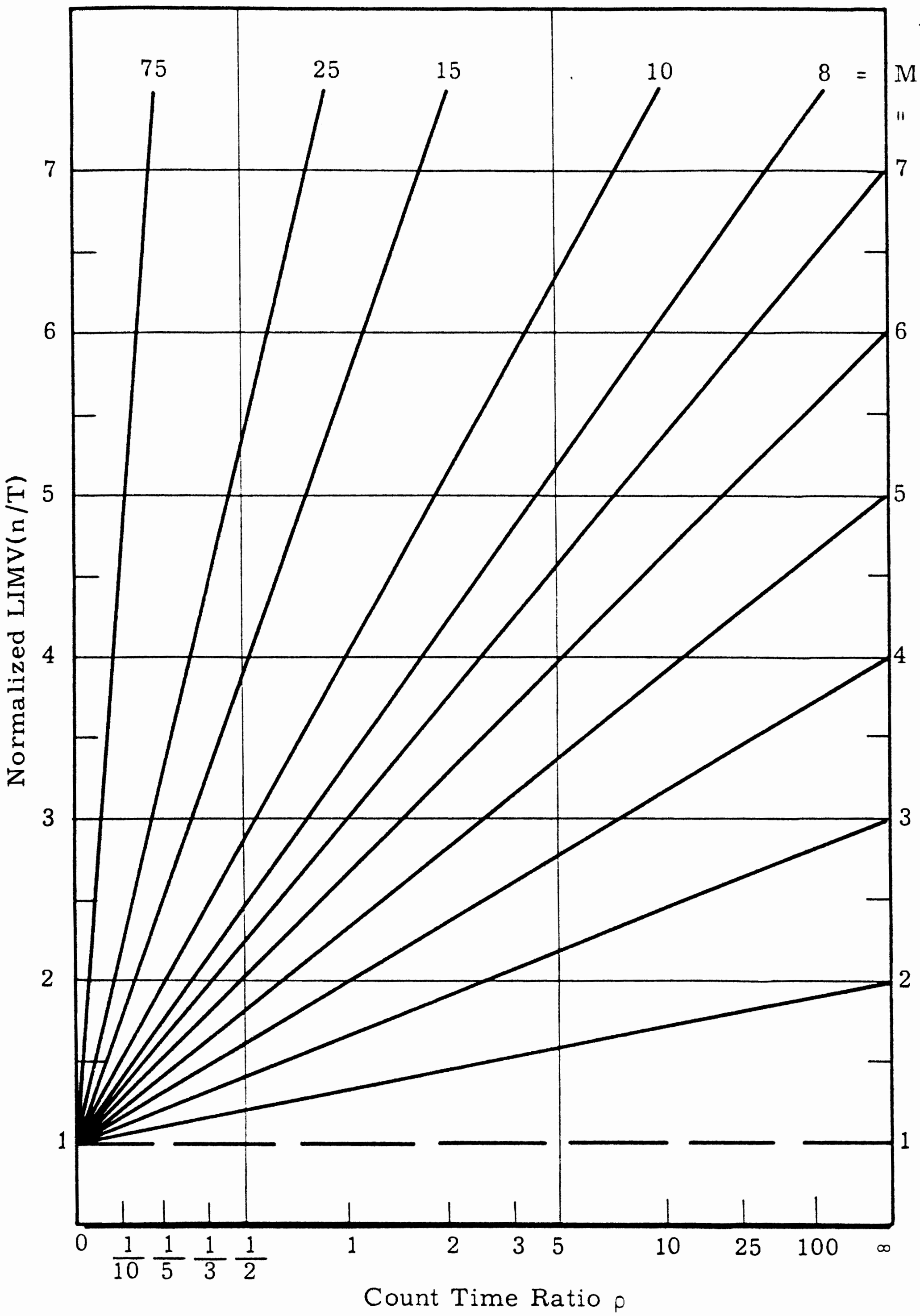

FIGURE 12

Normalized LIMV(n/T) for 5% Level of Significance Decision Rule D_3 as a Function of Number M of Samples Corrected with Same Background and Count Time Ratio ρ

Let $\text{msd}\,(M, \rho) = [\underset{k}{\text{Max}}\ \text{LIMV}(n/T)]^{1/2}$; i.e., $\text{msd}\,(M, \rho)$ is the maximum asymptotic standard deviation of n/T for fixed M and ρ. Several useful benchmarks from Figure 12 are :

I. With $M \leqq 4$, $\text{msd}\,(M, \rho) \leqq 2\ \text{msd}\,(1, \rho)$ for all $\rho < +\infty$.

II. With $M \leqq 10$, $\text{msd}\,(M, \rho) \leqq 2\ \text{msd}\,(1, \rho)$ for $\rho \leqq 1$.

III. With $M \leqq 75$, $\text{msd}\,(M, \rho) \leqq 2\ \text{msd}\,(1, \rho \leqq 1/10$.

VII. ACKNOWLEDGMENTS

The author is indebted to numerous scientists in the Hanford Laboratories for raising questions concerning the probabilistic properties of count rate estimates - questions which collectively provided the impetus for the above investigation. During the development and writing of the paper, discussion with and constructive criticism by C. A. Bennett were very helpful. W. J. Dixon also offered several pertinent suggestions.

D. E. Hudson wrote and debugged the FORTRAN language programs for all the digital computer calculations. The actual computations were done on the Hanford IBM 7090 operated by Electronic Data Processing personnel. Spot checks of IBM 7090 output and other desk calculator computations were done by Pauline Z. Stewart.

REFERENCES

1. J. O. Irvin. "The Frequency Distribution of the Difference Between Two Independent Variates Following the Same Poisson Distribution," J. Royal Statistical Society, ser. B, vol. 100, p. 415. 1937.
2. J. Strackee, and J. J. Deiner van der Gon. "The Frequency Distribution of the Difference Between Two Poisson Variates," Statistica Neerlandica, vol. 16, pp. 17-23. 1962.
3. J. Neyman, and E. S. Pearson. "On the Problem of the Most Efficient Tests of Statistical Hypotheses," Phil. Trans. Roy. Soc., London, ser. A, vol. 231, pp. 289-337. 1933.
4. A. Birnbaum. "Statistical Methods for Poisson Processes and Exponential Populations," J. Amer. Stat. Assoc., vol. 49, pp. 254-266. 1954.
5. Conference on Measurements and Standards of Radioactivity, National Academy of Sciences - National Research Council, Washington, D. C., 1958. pub. 573, p. 110.
6. H. Cramer. Mathematical Methods of Statistics, Princeton University Press, Princeton, N. J., 1946.
7. W. J. Dixon, and F. J. Massey. Introduction to Statistical Analysis, McGraw-Hill Book Co., Inc., New York, N. Y., 1957.
8. S. S. Wilks. Mathematical Statistics, John Wiley & Sons, Inc., New York, N. Y., 1962.
9. J. Neyman. "On Statistics the Distribution of Which is Independent of the Parameters Involved in the Original Probability Law of the Observed Variables," Stat. Res. Mem., vol. 2, p. 58. 1938.
10. D. G. Chapman. "On Tests and Estimates for the Ratio of Poisson Means," Annals. Inst. Stat. Math., Tokyo, vol. IV, no. 1, pp. 45-49. 1952.
11. W. Feller. An Introduction to Probability Theory and Its Applications, John Wiley & Sons, Inc., New York, N. Y., 1961.
12. D. Durand. "Joint Confidence Regions for Multiple Regression Coefficients," J. Amer. Stat. Assoc., vol. 49, pp. 130-146. 1954.
13. J. Mandel. "A Note on Confidence Intervals in Regression Problems," Annals Math. Stat., vol. 29, pp. 903-907. 1958.
14. M. H. Hansen, W. N. Hurwitz, and W. G. Madow. Sample Survey Methods and Theory, Volume II - Theory, John Wiley & Sons, Inc., New York, N. Y., 1953.
15. Table of the Bivariate Normal Distribution Function and Related Functions, National Bureau of Standards Applied Mathematics Series, No. 50, Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C., 1959.
16. G. Friedlander, and J. W. Kennedy. Nuclear and Radiochemistry, John Wiley & Sons, Inc., New York, N. Y., 1955.